The reaction mechanism and inhibition of ATP-PRTase enzymes

Gerard Johan Moggré
(MSc)

A thesis submitted in partial fulfilment
of the requirements for the degree of
Doctor of Philosophy in Chemistry at the
University of Canterbury

Ferrier Research Institute

February 2016

Supervisor: Prof. Emily Parker, University of Canterbury
Co-supervisor: Prof. Peter Tyler, Ferrier Research institute

Abstract

Adenosine triphosphate phosphoribosyltransferase (ATP-PRTase) catalyses the first committed step in histidine biosynthesis. This thesis discusses studies towards a better understanding of the enzyme structure, function and inhibition by both natural and non-natural ligands, towards the goal of drug development. ATP-PRTase from *Campylobacter jejuni* and *Mycobacterium tuberculosis*, which both carry the long ATP-PRTase enzyme, and from *Lactococcus lactis* which carries the short form ATP-PRTase enzyme were studied using a range of approaches and techniques.

The structural basis of inhibition of the *C. jejuni* ATP-PRTase was investigated. An altered model of inhibition is proposed, based on two crystal structures: an active ATP-bound structure, and an inactive histidine and adenosine monophosphate (AMP)-bound structure.

Intrinsic kinetic isotope effect (KIE) values and transition state modelling was performed for all three ATP-PRTase enzymes. The reaction was studied in reverse, using phosphonoacetic acid as an alternative substrate to pyrophosphate to overcome the suppression of KIEs by commitment to catalysis. Radioactive labels were incorporated by coupled enzymatic syntheses at six positions, with the intrinsic KIEs measured using the internal competition method. A proposed dissociative reaction mechanism was supported by moderate primary ^{14}C isotope effects (1.028 to 1.031) and large α-secondary ^{3}H isotope effects (1.147 to 1.250). The transition state models were located with density functional theory at the B3LYP/6-31*G(d,p) level, and refined the reaction mechanism to a $D_N{*}A_N^{\ddagger}$ type.

Synthesis and inhibition studies were also performed, using transition state analogue synthesis, and preliminary work with a fragment based lead discovery technique. Three analogues were developed, with one based on the ring flattening effect predicted for the carbocation intermediate (a cyclopentene based analogue), and two based around the charge change of the cationic species (iminoribitol based analogues). The cyclopentene analogue was found have an IC 50 of 48 ± 4 μM for the *M. tuberculosis* ATP-PRTase enzyme and a K_i of 282 ± 66 μM for *C. jejuni* ATP-PRTase.

The combined results contribute to the understanding of the reaction mechanisms and inhibition of ATP-PRTase enzymes. The results present the first crystallographic observation

of ATP bound to the ATP-PRTase enzyme. KIE and transition state modelling of the reaction of Type IV PRTases was successfully demonstrated. Competitive inhibition was observed for an initial transition state analogue.

Acknowledgements

This thesis would not be here without help and support from many people. First and foremost I would like to thank Professor Emily Parker. Thank you for introducing me to the challenging but exciting world of biochemistry. I have been very grateful for this opportunity to work with you. The journey of my PhD research would not have been possible without your continuing support and hard work. Your passion and dedication to science is inspiring and is something I hope to keep with me for whatever the future will bring.

My thanks also go to my co-supervisor Professor Peter Tyler for his continuing support during my PhD. I value the wealth of knowledge you provided during the different stages of the project and your ever ready knowledge of anything synthesis related.

The ATP-PRTase project would not have come into existence without the pioneering efforts of Dr Gerd Mittlestädt. Thank you also for showing me the ropes in the biochemistry laboratory and teaching me the little tricks of the trade. I have greatly enjoyed working together on the ATP-PRTase project, and without your help and experience this project would not have been possible.

To Dr Logan Heyes, thank you for making the chemistry lab a fun place to work. I value your ability to critically discuss results and general insightful discussions both science related and not. Thank you also for the numerus “necessary” coffee breaks and the good times at conferences and the Australian synchrotron.

The Parker lab would not have been the same without the present and past group members. Thank you for making the lab a pleasant place to work. Special thanks goes out to Penel and Nicky for their support in the lab and insightful discussions and to Tom, Steve, and Yu for their roles in keeping the coffee flowing.

A great thank you also to Professor Vern Schramm at Albert Einstein College of Medicine for giving me the opportunity to come to New York City to gain first-hand experience in the art of kinetic isotope effect measurements and transition state modelling. The passion you show for science is inspiring.

A big thank you also to the group members of the Schramm laboratory, for making me feel welcome and letting me use all of your equipment. In particular, the work would not have

been possible without the support from Dr Myles Poulin, thank you for answering my many questions and teaching me the finer details of KIE measurements.

Thank you, Scott Cameron, Logan Heyes, Sarah Kessans and Alicia Moggré for carefully reading my thesis and providing helpful comments to improve the work.

To my family, thank you for the support you give me. In particularly, this thesis would not have been here without the continuing support of my wife Alicia, thank you for your encouragement throughout my research, for motivating me when science was hard and sharing the successes along the way. You help bring out the best in me. Aan mijn vader en moeder, Jan en Tini, en mijn broer Jan-Willem, bedankt voor jullie aanmoediging in alles wat ik doe zelfs als dat aan de andere kant van de wereld is.

Table of contents

Abbreviations

6PG	6-phosphogluconate
ADP	Adenosine diphosphate
AICAR	5-Aminoimidazole-4-carboxamide ribonucleotide
AK	adenylate kinase
AMP	Adenosine monophosphate
An-PRTase	Anthranilate phosphoribosyltransferase
ATP	Adenosine triphosphate
ATP-PRTase	Adenosine triphosphate phosphoribosyltransferase
Boc	*tert*-butyloxycarbonyl
CPM	Counts per minute
CSS	complexation significance score
DBMP	2,6-di-*tert*-butyl-4-methylpyridine
DFT	density functional theory
DMAP	4-(dimethylamino)pyridine
DMSO	dimethyl sulphoxide
DNA	deoxyribonucleic acid
DSF	differential scanning fluorimetry
EDTA	Ethylenediaminetetraacetic acid
EIE	Equilibrium isotope effect
EMA	European Medicines Agency
FBLD	Fragment based lead discovery
FDA	United States Food and Drug Administration
G6P	Glucose 6-phosphate
Glu	glutamic acid
GS	Ground state
GST	glutathione sulfur-transferase
HF	Hartree-Fock
HGPRTase	hypoxanthine-guanine phosphoribosyl-transferase
HIV	human immunodeficiency virus
HTS	high throughput screening
HWE	Horner-Wadsworth-Emmons
IAP	imidazole acetolphosphate
IGP	D-*erythro*-imidazole glycerol phosphate
IMAC	immobilised metal affinity chromatography
IPTG	isopropyl β-D-1-thiogalactopyranoside
IR	infrared
IUPAC	International Union of Pure and Applied Chemistry
KIE	Kinetic isotope effect
LB	lysogeny broth
MBP	Methyl bisphosphonate
m-CPBA	*meta*-chloroperoxybenzoic acid

MDR	multidrug resistant
MEPS	Molecular electrostatic potential surface
MsCl	methanesulphonyl chloride
MTAN	5'-methylthioadenosine nucleosidase
$NADP^+$	nicotinamide adenine dinucleotide phosphate
NADPH	Reduced nicotinamide adenine dinucleotide phosphate
n-BuLi	*n*-butyllithium
NMR	Nuclear magnetic resonance
OMP	orotidine monophosphate
OPRTase	Orotate phopshoribosyltransferase
PA	Phosphonoacetic acid
PCR	polymerase chain reaction
PDB	Protein Data Bank
PEP	phosphoenol pyruvate
PES	potential energy surface
PES	potential energy surface
PK	pyruvate kinase
PNP	purine nucleoside phosphorylase
PPase	pyrophosphatase
PP_i	Pyrophosphate
PRAMP	phosphoribosyl-adenosine monophosphate
PRATP	Phosphoribosyl adenosine triphosphate
PRPA	phosphoribose phosphonoacetic acid
PRPP	Phosphoribosyl pyrophosphate
PRPPase	phosphoribosyl pyrophosphate synthetase
PRTase	phosphoribosyltransferase
PYR	pyruvate
R5P	ribose-5-phosphate
RCM	Ring closing metathesis
RMSD	root-mean-square deviation
Ru5P	ribulose-5-phosphate
SAR	structure activity relationship
SEC	size exclusion chromatography
SEM	standard error of the mean
SPE	solid phase extraction
STD-NMR	Saturation transfer diffusion NMR
TAE	Triethyl amine
TB	Tuberculosis
TBDMSCl	*tert*-butyldimethylsilyl chloride
TEV	tobacco etch virus
Tf	trifluoromethanesulphonyl
THF	Tetrahydrofuran
THP	tetrahydropyran

Tris	tris(hydroxymethyl)aminomethane
TS	Transition state
TST	transition state theory
XDR	extensively drug resistant
ZPE	zero-point energy
αKG	α-ketoglutarate

List of Figures

List of Tables

Chapter 1

Introduction

The work presented in this thesis is focused on the first committed step in the histidine biosynthetic pathway. Efforts towards a better understanding of the enzyme structure, function and inhibition through both natural and non-natural ligands are presented. To build a foundation for understanding these processes, this chapter introduces the target organisms, followed by histidine biosynthesis, the family relationship of the target enzyme, and the structure and inhibition of the enzymes of interest.

1.1 Enzymes

Enzymes are biological catalysts that have the ability to specifically enhance reaction rates, impose stereochemical control, and provide means of regulation of complex metabolic processes.[1] They catalyse an enormous diversity of chemical reactions, with each enzyme specifically binding a small number of substrates and generally catalysing only one reaction. Rate enhancements of up to 10^{15} times that of the uncatalysed reaction are achieved under dilute and mild conditions.[1,2] Placed in sequence, enzymes form biosynthetic pathways that play vital roles in digestion, energy production and the biosynthesis of the complex metabolites essential to all living organisms. The step at the start of a specific biosynthetic pathway is usually referred to as the committed step, as from this point a metabolite is committed to be converted to the end product of the pathway. Regulation, another key function of enzymes, often occurs at this step, controlling the flux into the pathway. Direct regulation of enzyme activity can be achieved by the reaction product subsequently binding to the enzyme or by the end product of the pathway. The slowing down of the catalytic rate by enzyme inhibition allows an organism to control the production of metabolites based on the cellular needs.

The ability to modulate the function of an enzyme through inhibition also forms the basis of many drugs.[3,4] Enzyme inhibitors can be used to correct a metabolic imbalance or to block an essential metabolic pathway, which can kill a pathogen. In this, it is the specificity of an enzyme that allows for its selective targeting without impacting other vital processes.

1.2 Target organisms

1.2.1 Tuberculosis

Tuberculosis (TB), which is estimated to affects 2 billion individuals worldwide, is predominately caused by an infection with *Mycobacterium tuberculosis*.[5] The organism was

first discovered in 1882 by Robert Koch and is a rod shaped bacterium.[6] *M. tuberculosis* has a unique lipid rich cell wall structure, which gives rise to the organism's pathogenicity and resistance to treatments. TB infections are generally found in the respiratory system, but infections of other parts of the body such as the spine and lymph nodes can occur.[7,8] In pulmonary TB, the infection triggers an immune response which leads to the formation of a macrophage-rich cell mass known as granuloma or tubercles, after which the disease is named.[9,10] Formation of granuloma prevents the spreading of the disease to the remainder of the lung and other organs; however it does not effectively kill the bacteria. The encapsulated bacteria can survive in these granuloma for extended periods, which is referred to as a latent TB infection.[11] TB is a widespread disease affecting around one third of the world's population, of which most infections are latent.[12–14] Individuals infected with a latent form of the disease do not show symptoms and are not at risk of spreading the disease, but they do have a 5% risk of developing the active form of TB at some point in their lives.[12] Immune compromised individuals, such as those co-infected with human immunodeficiency virus (HIV), have an increased 50% risk of developing the active form of the disease.[12] The World Health Organisation millennium development goal target to halt and begin to reverse tuberculosis spread globally was reached in 2004.[14] However, the decline of the disease remains less than 1% per year, with over 9 million people still reported to fall ill to the disease each year.[13,15,16] In 2013 alone, the total number of deaths caused by TB was estimated to be 1.5 million people, with a quarter of these individuals co-infected by HIV.[17] The burden of 80% of the world's cases of the disease lies in only 22 countries, which are often located in poorer and less developed regions.[13] In these countries, the lack of appropriate diagnostics and limited treatment availability are the major causes of the prevalence of the disease. The standard treatment involves a combination of four frontline drugs isoniazid, a rifamycin, pyrazinamide and ethambutol, administered over a six to nine month period (Figure 1.1).[12]

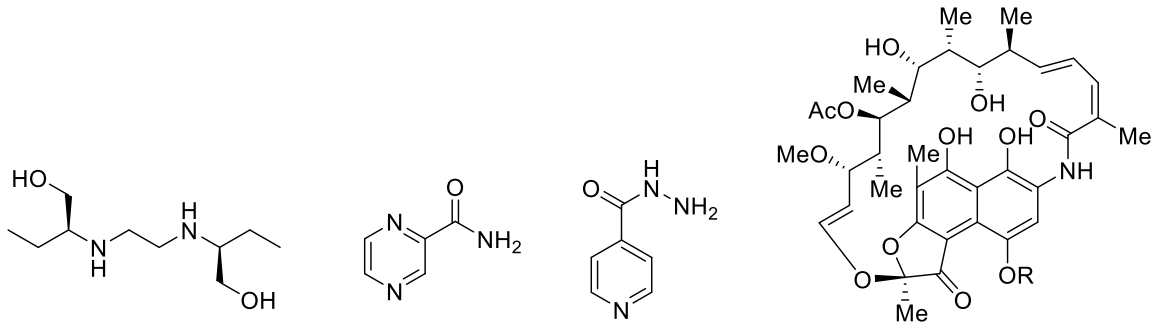


Figure 1.1. The four frontline drugs ethambutol, pyrazinamide, isoniazid and rifamycin (part of the ansamycins family) used for the treatment of TB-infected individuals. The general treatment regimen consists of a combination of these drugs over a 6-9 month period.

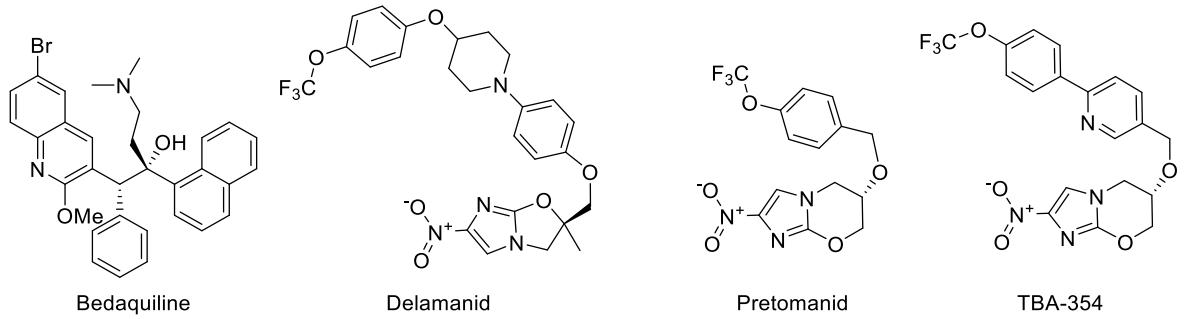
Although the available treatment regimens can cure most patients and effectively reduce transmission, they are not without problems. The duration and complexity of the combination of drugs can lead to failure and relapse and the development of drug resistance. Side effects in response to the anti-TB drugs are common and are an additional contributing factor in the non-adherence of patients to their treatment regimen. Of particular concern is increasing prevalence of TB which is resistant to the available treatments. Based on this resistance, TB can be divided into two groups: multidrug resistant (MDR) strains and extensively drug resistant (XDR) strains. MDR-TB is non-responsive to at least isoniazid and rifampicin, whilst XDR-TB is unresponsive to rifampicin and isoniazid, as well as second line drugs such as fluoroquinolone, capreomycin, kanamycin, and amikacin.[14,15,18]

1.2.1.1 New drugs and drug regimens to treat TB

The emergence of MDR strains of *M. tuberculosis* and the lack of development of a new frontline drug over the last 50 years makes the discovery of new approaches for TB treatment based on new molecular scaffolds a priority.[12] Ideally these new treatments would be effective against MDR and XDR strains, have shorter treatment times, require low dose frequency, simple administration and be co-administrable with HIV medication.[12,19] The growing awareness around TB in recent years has led to an increase in research into TB drugs, with considerable progress being made over the last ten years. There are six new drugs in preclinical development, and eight drugs in phase I, II or III of clinical trials for the treatment of both drug-resistant and susceptible TB in its active or latent forms.[5] The new treatment regimen studies include the efficacy of existing firstline drugs applied in new

combinations, as well as increased doses to shorten treatment time. For example, high-dose daily rifapentine-containing regimens have been shown to reduce the intensive phase of treatment to less than six months. Fluoroquinolone-based antibiotics have also been studied, however a recent four month phase III clinical trial gave disappointing results in which the substitution of ethambutol or isoniazid was associated with a higher risk of relapse of TB.[5]

A new class of compounds are the nitroimidazole-containing compounds such as TBA-354, delamanid and pretomanid (Figure 1.2).[5] These are pro-drugs, and require bioreduction of the nitrogroup for activation, catalysed by a deazaflavin-dependent nitroreductase.[20,21] Bioreduction leads to the release of reactive nitrogen species, which include nitric oxide, and is partly responsible for the anti-TB activity. TBA-354, which has just entered phase I clinical trials, was developed to further optimise potency of delamanid and pretomanid.[21] Delamanid is in phase III clinical trials and works by blocking the synthesis of mycolic acid, an important part of the *M. tuberculosis* cell wall.[22,23] Delamanid was granted accelerated approval by the European Medicines Agency (EMA) in 2013 for the treatment of XDR-TB.[5] This is the second drug receiving accelerated approval following bedaquiline in 2012, which was approved by the United States Food and Drug Administration (FDA) for the treatment of XDR-TB as a last resort drug (Figure 1.2).[24] Bedaquiline affects the proton pump of ATP synthase which is a critical enzyme in the synthesis of ATP for *M. tuberculosis*.[24] The drug binds 20,000 times more strongly to the *Mycobacterium* variant of this enzyme than to the human equivalent. This drug however gave rise to health concerns as it was shown to cause arrhythmia by prolonging the hearts QT time, which can lead to cardiac arrest.[24]



Bedaquiline Delamanid Pretomanid TBA-354

Figure 1.2. New TB drugs currently in clinical trials. Bedaquiline is part of diarylquinolines, a novel group of compounds for the treatment of TB. Delamanid, pretomanid and TBA-354 are all nitroimidazole-containing compounds and are in various stages of clinical trials.

1.2.2 *Campylobacter jejuni*

Campylobacter jejuni is a Gram-negative, spiral-shaped, non-spore forming bacterium. *C. jejuni* is motile through the presence of unipolar or bipolar flagella and is relatively slow growing.[25] The bacteria are one of the most common cause of gastroenteritis, and symptoms range from diarrhoea and abdominal cramps and pain to joint pain and fever in some cases.[25] Ingestion through contaminated food or water, uncooked poultry, and raw dairy products are the main causes of a *Campylobacter* infection. An infection normally shows symptoms 3–5 days after consumption and can take somewhere from 1 hours to 7 days to clear in most cases. The infection has been linked to a form of neuromuscular paralysis known as Guillain-Barre syndrome, an autoimmune disease affecting the nervous system. This syndrome develops after the infection, but cases are rare.[26] A *Campylobacter* infection can be severely debilitating, but generally does not require treatment with antibiotics. Treatment through liquid and electrolyte replacement is commonly sufficient to clear the disease.[27] However, in cases involving prolonged illness, high fever, bloody stools, pregnancy, or co-infection with HIV, treatment of the infection with antibiotics can become necessary.[25,27,28] When antibiotic treatment is required, macrolides such as clarithromycin or azithromycin are the main form of treatment for confirmed *Campylobacter* cases. Fluoroquinolones are also commonly used, as *C. jejuni* infections cannot be distinguished from diarrhoeal illnesses, caused by *Salmonella* and *Shigella*, based on symptoms alone (Figure 1.3).[25] Alternative drugs that can be used to treat a *C. jejuni* infection are the tetracycline doxycycline, and chloramphenicol (Figure 1.3). However, the use of antibiotics in the food industry has led to the development of significant resistance with resistance observed up to 57% for tetracycline based antibiotics in some cases.[25,29] *Campylobacter* remains a significant cause of food borne illnesses and continues to be a burden on public health.

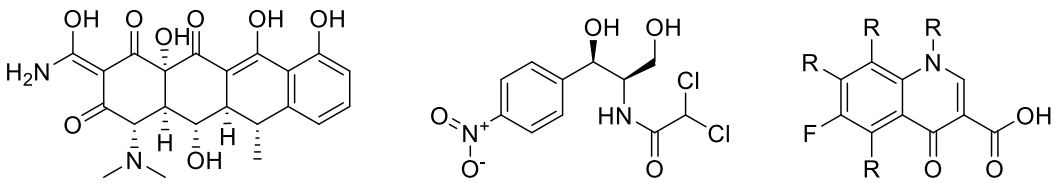


Figure 1.3. Common antibiotics used for the treatment of a C. jejuni infection; *doxycycline, chloramphenicol and the essential structure for fluoroquinolones.*

1.3 Amino acid biosynthesis as a drug target

There are 22 proteinogenic α-amino acids from which proteins are synthesised. Thirteen of these can be synthesised by humans, making the remaining nine, which include L-histidine, essential to our diet to ensure optimal growth and development.[1] Bacteria, fungi and plants, on the other hand, do have the biosynthetic pathways for these amino acids. The absence of these pathways in humans makes enzymes of these biosynthetic pathways potentially interesting for the design of antimicrobial drugs and fungicides which are non-toxic to humans.[4,30] For example, inhibition of histidinol dehydrogenase, involved in histidine biosynthesis, showed growth inhibition for *Brucella suis* in minimal media and human macrophages.[31] Also, *M. tuberculosis* was shown to be auxotrophic for L-histidine or L-tryptophan when enzymes involved in the biosynthesis of these amino acids were deleted.[32] This study further showed that these L-tryptophan or L-histidine auxotrophs did not survive single amino acid starvation experiments, which is similar to the environment encountered in the compartment of macrophages where the TB bacteria reside.[33,34] This supports the theory that enzymes of this pathway are potential drug targets for the treatment of TB and other infectious diseases.[32]

1.4 Histidine biosynthesis

The biosynthesis of L-Histidine, referred to as histidine from here on, comprises ten linear steps (Figure 1.4).[35] This pathway is energetically expensive and to control production the pathway is regulated both at the enzyme level and the transcription level.[36] The first committed step involves the condensation reaction between adenosine triphosphate (ATP) and phosphoribosyl pyrophosphate (PRPP) to form phosphoribosyl-ATP (PRATP) and inorganic pyrophosphate (PP_i), catalysed by adenosine triphosphate phosphoribosyltransferase (ATP-PRTase) (Figure 1.4).[37] As is commonly observed for enzymes involving the first committed step of a biosynthetic pathway, the activity is subject to feedback inhibition. The PRATP product then undergoes another four steps to form 5-aminoimidazole-4-carboxamide ribonucleotide (AICAR) and D-erythro-imidazole-glycerol-phosphate (IGP). This is where the pathway branches. AICAR feeds into the purine biosynthetic pathway, whilst IGP undergoes a further five steps to give the end product histidine.[35]

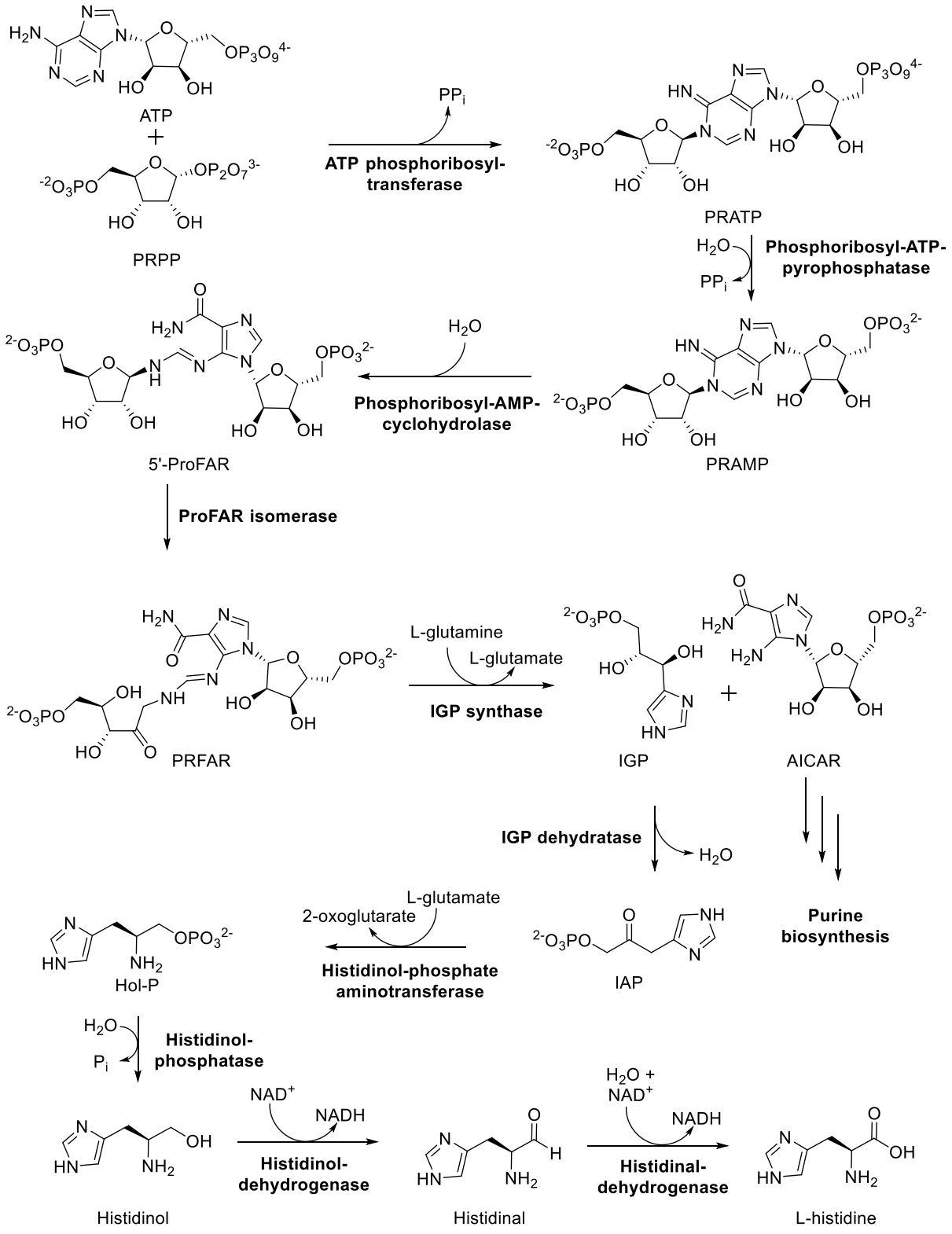


Figure 1.4. Schematic overview of histidine biosynthesis.

1.5 Adenosine triphosphate phosphoribosyltransferase

Adenosine triphosphate phosphoribosyltransferase is the enzyme responsible for the first committed step in histidine biosynthesis and the focus of this thesis. ATP-PRTase catalyses the reversible, magnesium-dependent, substitution reaction between ATP and PRPP (Figure 1.5).[37] ATP-PRTase controls the flux of metabolites into the histidine biosynthetic pathway through an allosteric feedback mechanism by the product of the pathway, histidine, providing control of the enzyme as a function of the available histidine. Adenosine monophosphate (AMP) and adenosine diphosphate (ADP) act as competitive inhibitors and therefore reflect on the energy state of the cell. In some cases higher concentrations of the PRATP product also inhibits activity.[37]

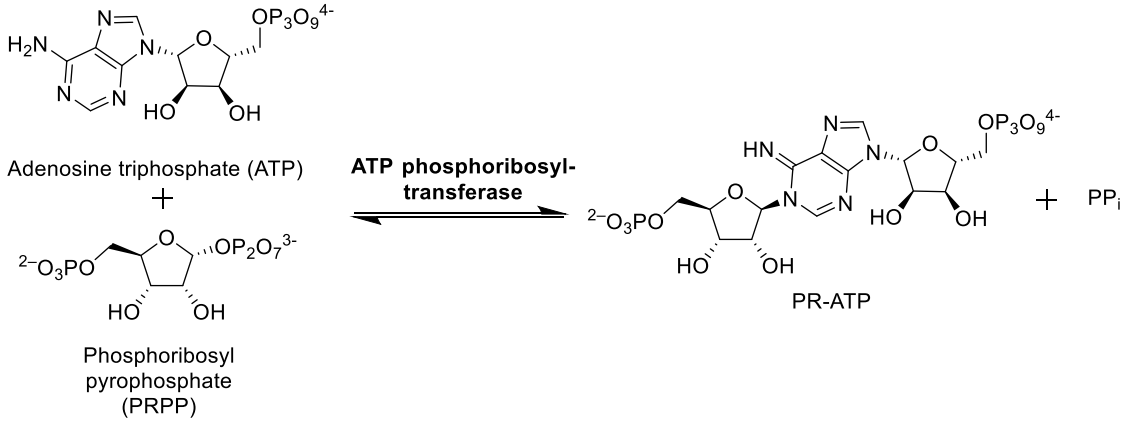


Figure 1.5. The ATP-PRTase enzyme catalyses the condensation reaction between PRPP and ATP to form PRATP and PP_i.

1.6 Classification of the PRTase enzyme family

ATP-PRTase belongs to the phosphoribosyltransferase (PRTase) enzyme family. This group of enzymes generally catalyse the transfer of the ribose 5-phosphate part of PRPP to a nitrogenous, and generally aromatic, base.[38] However, not all PRTase proteins perform biosynthetic roles. In some bacteria PRTase enzymes are also involved in regulation of the expression of purine and pyrimidine synthetic genes. The proteins of this family play diverse metabolic roles such as purine and pyrimidine biosynthesis, as well as in the nucleotide salvage pathway.[39–42]

Some of the best-studied types of PRTase enzymes are the Type I PRTases and these were first identified at the sequence level.[38] Crystal structures later confirmed the family relationship of the Type I PRTases.[38] The Type I PRTase enzymes share a common core responsible for PRPP binding; additional decorations, and in some cases an additional domain is responsible for binding the nucleophile. The common core consists of a five stranded parallel β sheet (β2-β1-β3-β4-β5) decorated by three α helices. α1 and α2 are found on one face of the β sheet and α3 is found on the other side (Figure 1.6, **A**). There is a thirteen amino acid sequence motif characterised by four hydrophobic amino acids, two acidic amino acids of either glutamate or aspartate and seven amino acids of variable character, usually including glycine and threonine.[38] This motif makes up the PRPP binding site which in itself is made up of three loops, a PP_i loop; a flexible loop and a phosphate loop. The phosphate loop constitutes nine residues and is responsible for binding the ribose 5-phosphate parts of PRPP. The two acidic amino acids, which are generally found adjacent to each other at the start of the phosphate loop, are responsible for binding the hydroxyl groups of the ribose ring. The PP_i binding loop is a small loop that accommodates the pyrophosphate part of PRPP and binding of this ligand is between the phosphate and PP_i binding loops. The flexible loop acts as a gate and sits over the active site. PRPP binds to this signature sequence in association with magnesium.

The classification of a second type of PRTase enzymes was proposed based on the novel fold found for quinolinic acid phosphoribosyltransferase (Figure 1.6, **B**).[43,44] The Type II PRTases lack the signature sequence and the classification of this enzyme as part of the PRTase family was based on the reaction chemistry. Nicotinic acid phosphoribosyltransferase shares this unique fold and together with quinolinic acid phophoribosyltransferase these two enzymes form the Type II subgroup. The active site architecture is conserved among bacteria and eukaryotes for this Type. Differences are found in the quaternary structure with dimeric and hexameric species reported.[43,44] The novel fold consists of a mixed α/β *N*-terminal domain and a unique seven-stranded α/β barrel-like domain.[45] The PRPP binding site is formed at the interface of two chains with residues contributing from both chains. Given the vastly different active site architecture between Type I and Type II PRTases a remarkable similarity between the binding mode of PRPP is observed, with differences only found around the pyrophosphate part which is rotated approximately 60° compared to Type

I enzymes. Interestingly, two acidic amino acids shown to interact with the hydroxyl group of the ribose ring are also present in Type II enzymes. Because of its unique fold, Type II PRTase has most likely evolved independently of Type I PRTases.

Type III PRTase enzymes are exemplified by anthranilate PRTase (An-PRTase), which catalyses the second step in tryptophan biosynthesis (Figure 1.6, **C**).[46] An-PRTase is a homodimer, with the active site found in the hinge region of the two domains that make up the monomeric unit. A large C-terminal domain formed by eight α-helices surround a central seven-stranded β-sheet. The *N*-terminal domain, which is significantly smaller, is made up of six α-helices. PRPP binding is mostly accommodated by the C-terminal domain. The binding mode of PRPP resembles that of the other types of PRTase enzymes with again differences mainly found around the pyrophosphate part. Again binding of PRPP is associated with a divalent metal ion essential for catalysis.

The ATP-PRTase enzyme belongs to Type IV subgroup of PRTases and like An-PRTase solely makes up this subgroup in the PRTase enzyme family (Figure 1.6, **D**).[47] ATP-PRTase is found as two structurally diverse enzymes which differ in their allosteric machinery. The long form is a homohexamer with the subunit comprised of a mixed α/β bilobal domain containing the active site, and a C-terminal ACT domain where the allosteric binding site for histidine is found. The short form is a heterooctamer, containing a similar mixed α/β bilobal catalytic core but instead of a regulatory domain, the short form enzyme associates with a second enzyme, structurally related to histidyl-tRNA-synthatases, to achieve allosteric regulation.[48] The active site is found in the cleft of domains I and II that make up the catalytic core. Domain I, a mixed α/β domain, makes up the ATP binding site and domain II shares parts of the signature amino acid sequence found for Type I PRTase enzymes and accommodates PRPP binding. Domain II has structural similarities to the shared core of the Type I PRTase subgroup, with a five stranded parallel β sheet decorated by three α helices. The PRPP binding site including the conserved acidic amino acids found in Type I PRTases including the phosphate and pyrophosphate binding loop. Type IV PRTases lack the flexible hood domain which sits over the PRPP binding site in Type I enzymes.

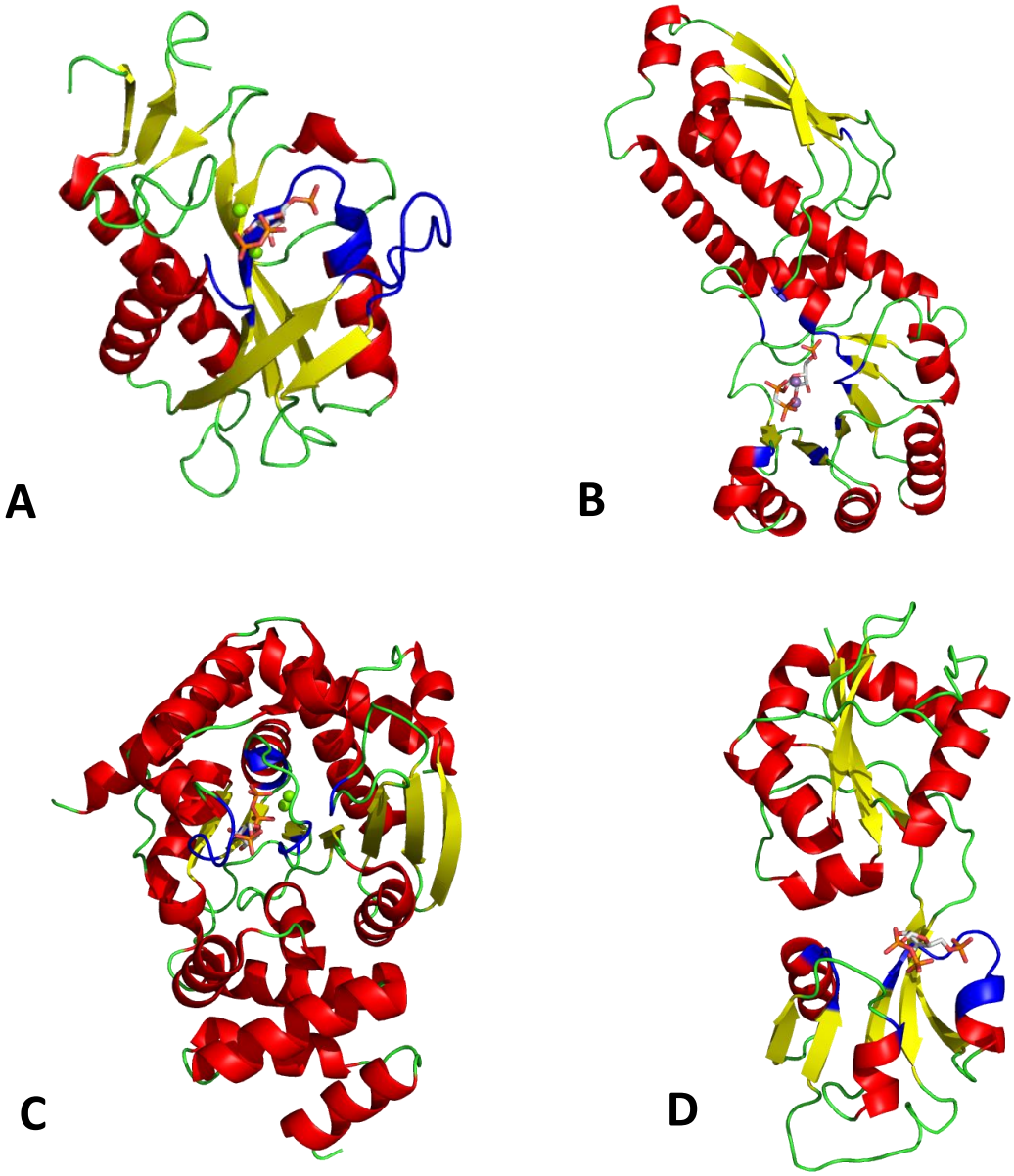


*Figure 1.6. The PRTase family can be divided into four types based on their structural architecture. Type I **A** Hypoxanthine phosphoribosyltransferase from Trypanosoma cruzi PDB:1TC2. Type II **B** Quinolinate phosphoribosyltransferase from M. tuberculosis PDB: 1QPR.Type III **C** anthranilate phosphoribosyltransferase from M. tuberculosis PDB:1ZX1. Type IV **D**, the shared catalytic core of ATP-PRTase is shown from Lactococcus lactis PDB: 1Z7N. α-helices are displayed in red, β-sheets in yellow, flexible loops in green, and active site regions involved in PRPP binding in blue.*

1.7 The structure of ATP-PRTase long form

The structure of ATP-PRTase long form comprises around 300 amino acids. Crystal structures for *Escherichia coli* (PDB: 1Q1K, 1H3D), *M. tuberculosis* (PDB: 1NH7, 1NH8) and *Methanobacterium thermoautotrophicum* (PDB: 2VD3), in both the unliganded or apo form

and ligand bound forms have been reported. The crystal structures revealed an elongated molecule consisting of 10 α-helices and 15 β-sheets (Figure 1.7). Each monomeric unit of the protein is formed by three domains: domain I consisting of a six stranded β-sheet (β1-β5, β11) wrapped by four α-helices (α1-α3, and α8); domain II is topographically inserted into domain I and comprises five stranded β-sheets (β6-β10) and four α-helices; and domain III features an antiparallel β-sheet topology with a pair of antiparallel alpha helices on one face.[47,49,50] Domain III is where the allosteric histidine binding site is found, approximately 40 Å away from the active site, and resembles an ACT regulatory domain found in a variety of proteins including enzymes involved purine and amino acid metabolism.[50] Work done by Kryvi *et al.* on the *E. coli* ATP-PRTase suggests it exists in a dynamic equilibrium between an active dimeric form and an inactive hexameric form at low concentrations of the enzyme.[47,51]

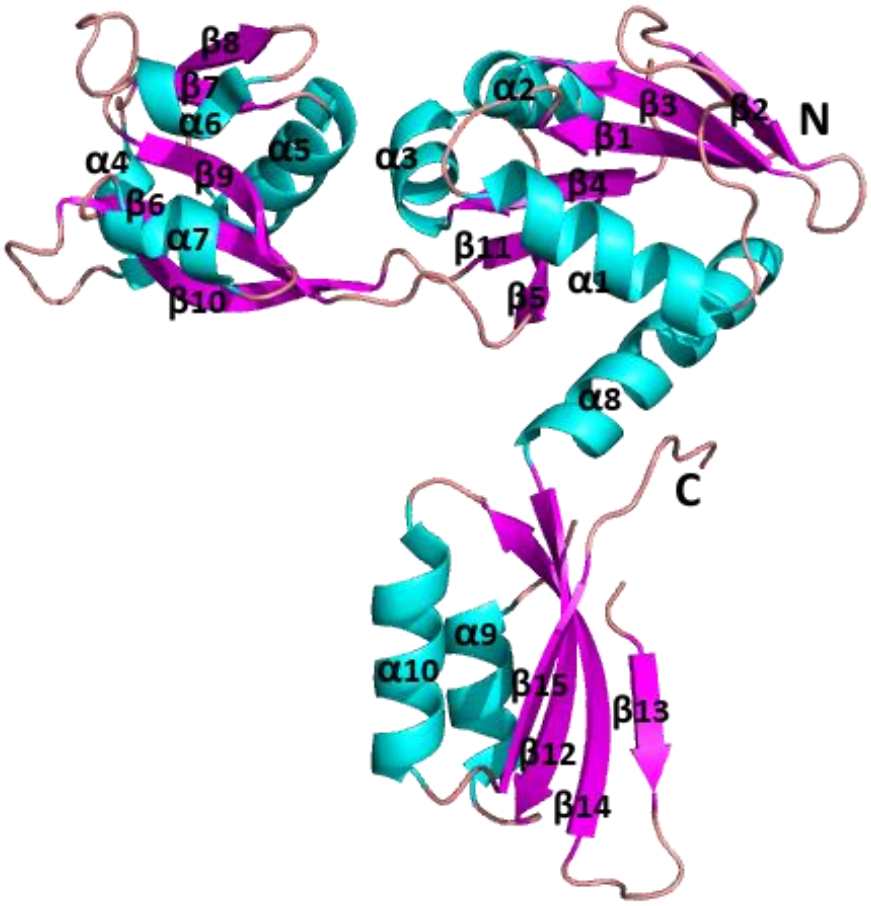


Figure 1.7. The monomeric unit of ATP-PRTase long form M. tuberculosis with histidine bound (PDB: 1NH8) displayed here with α-helices in cyan, β-sheets in purple and flexible loops in pink.

1.7.1 Natural allostery and inhibition

ATP-PRTase activity is controlled through different mechanisms, allosterically through a feedback mechanism by the end product histidine, synergistically with AMP and histidine, and inhibition by AMP or ADP to reflect the overall energy state of the cell.[37,52] High concentrations of ATP and high concentrations of the product PRATP sometimes also result in reduced activity.[51]

M. tuberculosis ATP-PRTase was the first long form of this enzyme successfully crystallised as the apo structure (PDB: 1NH7) and the synergistically inhibited, histidine and AMP bound, structure (PDB: 1NH8). Both structures were crystallised in the hexameric form, which was attributed to the high protein concentration required in the crystal conditions. The apo structure displayed a more open and loose hexamer then the inhibited enzyme. Crystal structures with AMP, *E. coli* ATP-PRTase (PDB: 1H3D), and histidine, *M. thermoautotrophicum* (PDB: 2VD3), have since also been reported. The AMP bound structure resembled the apo structure, where the histidine bound structure displayed the more closed conformation, suggesting that histidine alone is responsible for the observed structural change.

1.7.1.1 Histidine binding site and the regulatory response

The histidine binding site is found approximately 40 Å away from the active site and regulation is thus achieved allosterically. Histidine binding stabilises the closed conformation, which has an altered orientation of the regulatory domain relative to the catalytic core as compared to the monomer of the apo structure. The histidine binding site is found at the interface of two domains and involves residues from a flexible loop (α9–β13) and Leu253 on β-sheet 14 from one monomer, and of residues Ala273' (β15) and Asp218' (β12) of the adjacent monomer (Figure 1.8, **A**). The interaction of histidine across the interface of two domains brings the regulatory domains closer together. the cross interface interaction is believed to be the basis of stabilisation of the hexameric form of the enzyme (Figure 1.8, **B**).[49]

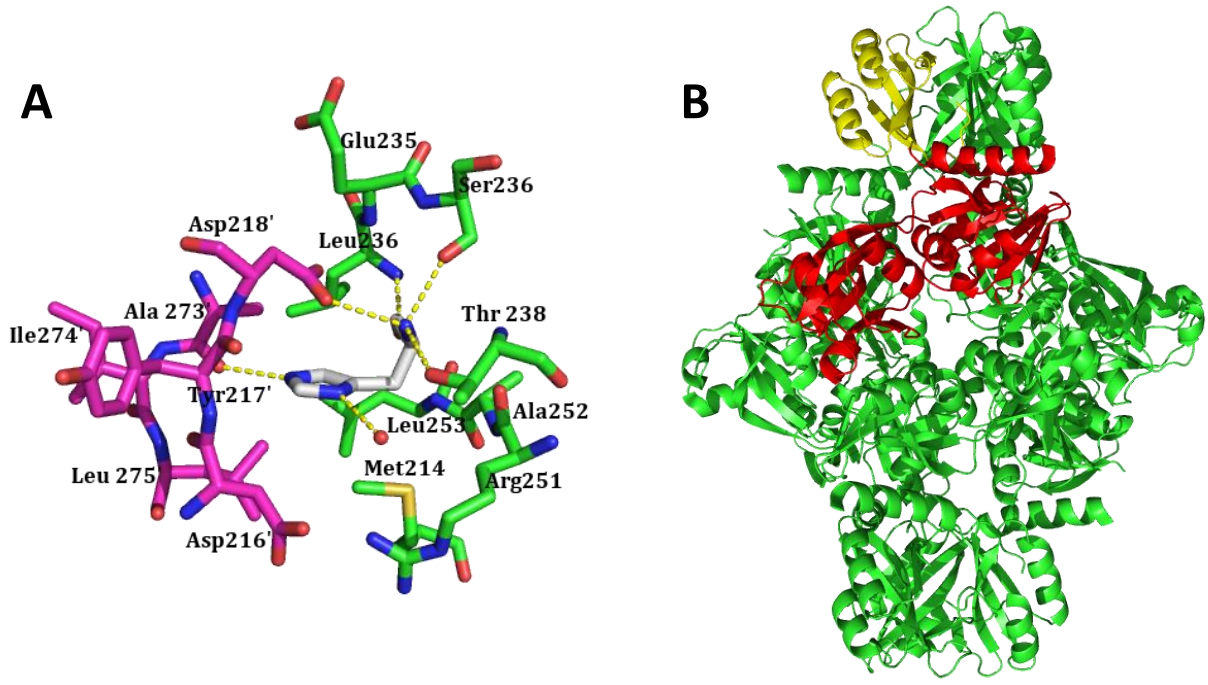


Figure 1.8. Histidine binding site of M. tuberculosis (PDB: 1NH8). ***A*** *Histidine (white) makes interactions across the interface of two regulatory domains with the main chain (green) and the adjacent monomer (pink). Upon binding of the allosteric ligand the protein undergoes a conformation change bringing the regulatory domains closer together. Heteroatoms are coloured according to element: oxygen (red), nitrogen (blue), phosphorous (orange).* ***B*** *Histidine stabilises the inhibited hexameric form of the ATP-PRTase long form enzyme (PDB: 1NH8). A single chain is coloured with the catalytic core (domain I and II) in red and the regulatory domain in yellow.*

1.7.1.2 Synergistic inhibition of ATP-PRTase by histidine and AMP

The mode of action for AMP inhibition was initially believed to be through competitive binding at the ATP site.[49] The crystal structures of *M. tuberculosis* and *E. coli* ATP-PRTase, however, reveal that binding of AMP occurs across the PRPP and ATP binding site (Figures 1.9 and 1.10).[47,49] Looking at the binding mode of AMP, insight into how AMP can act as a competitor for both ATP and PRPP, and how AMP provides a synergistic inhibitory response with histidine, can be elucidated.

In the AMP bound *M. tuberculosis* ATP-PRTase structure the 5-phosphate of AMP binds in the phosphate binding loop (Asp154 -Thr161) of the PRPP binding site (Figure 1.9). Here, a hydrogen bonding (H-bonding) network is formed between the phosphate group and Gly157–Thr161. The adenine ring sits over a hydrophobic part of the active site towards domain I. The 6-amino group of adenine forms a H-bond with Tyr116, and N1 interacts with

three ordered water molecules present in the ATP binding site. The 2′ and 3′ hydroxyl groups of ribose form interactions with Asp30′ across the dimer interface to the adjacent monomer. Formation of these interactions could explain further stabilisation of the inhibited form in the presence of AMP and histidine.[47] It should be noted that for the *M. tuberculosis* structure, the electron density for the 2′ and 3′ hydroxyl group and the CH_2 at the 5′ position of the ribose ring of AMP is of poor quality, with no density observed at these positions.[47]

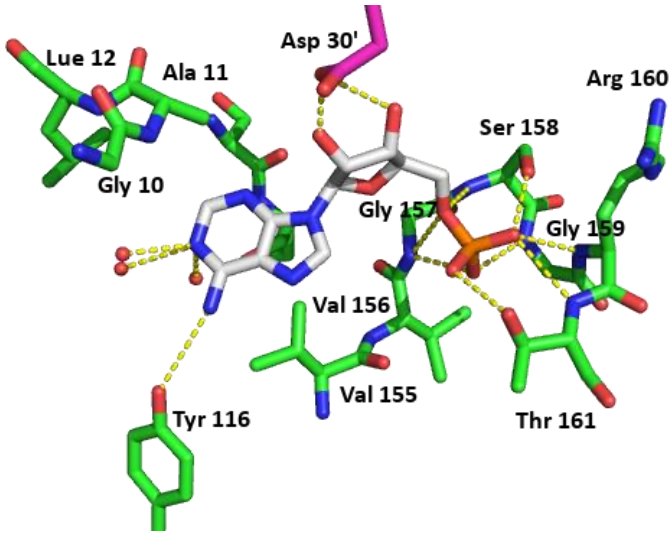


Figure 1.9. The binding mode of AMP to M. tuberculosis ATP-PRTase. AMP binds at the interface of two monomeric units forming interactions across the interface. The main chain residues are in green and residues from the adjacent monomer are in purple. Heteroatoms are coloured according to element: oxygen (red), nitrogen (blue), phosphorous (orange).

In the *E. coli* ATP-PRTase structure with AMP bound to the active site, a different binding mode is modelled for AMP (Figure 1.10). The ribose ring is 180^{o} rotated compared to its position in the *M. tuberculosis* structure and only interacts with a single chain. The difference in binding may be a result of the absence of Arg16 in the *M. tuberculosis* ATP-PRTase enzyme, providing extra stabilisation of the adenine part of AMP in the *E. coli* structure. The ribose 2′ and 3′ hydroxyls form additional interactions with the two acidic residues Asp169 and Glu156 placing this part of AMP in a similar position as PRPP.[49]

The ability of AMP to bind across both the PRPP and ATP binding sites explains the basis of the competitive nature of AMP with respect to both substrates.

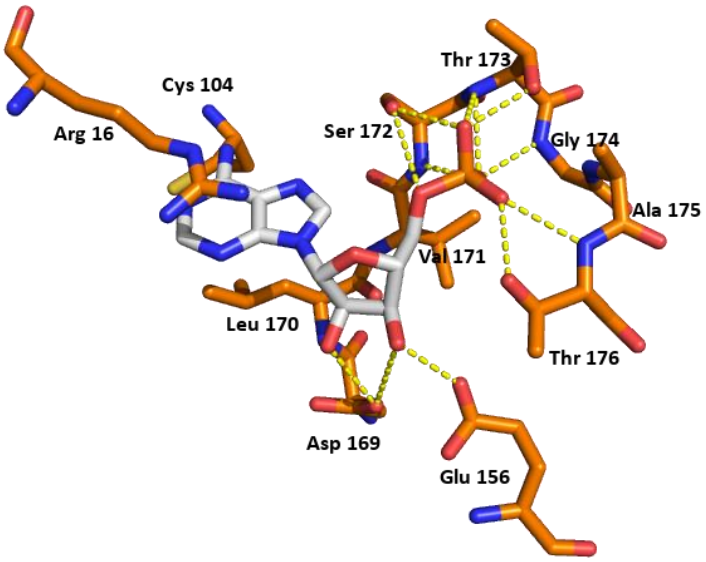


Figure 1.10. AMP (white) binding mode for E. coli ATP-PRTase forming most binding interactions with the PRPP binding site and additional stabilisation of a non-conserved Arg16 positioned over the active site. The main chain is coloured in brown. Heteroatoms are coloured according to element: oxygen (red), nitrogen (blue), phosphorous (orange).

1.8 The structure of ATP-PRTase short form

The short form ATP-PRTase comprises 205-220 residues and is structurally similar to domain I and II of the long form enzyme (Figure 1.11, **A**).[53,54] The ATP-PRTase short form lacks the 65–105 residues long C-terminal regulatory domain and instead associates with a second enzyme for regulation.[48] Sequence comparison revealed distant similarities between the additional unit and an aminoacyl-tRNA synthetase enzyme.[48] The ATP-PRTase short form enzyme comes together with the aminoacyl-tRNA synthetase-like domain to form a hetero-octamer containing four chains of each (Figure 1.11, **B**). The overall architecture features an X-shaped complex made up of the aminoacyl-tRNA synthetase units. The ATP-PRTase catalytic core is found as a dimer on each end. Crystal structures of the full complex have been obtained for *Thermotoga maritima* in complex with histidine and *L. lactis* with phosphate bound (PDB: 1Z7M) and PRPP bound (PDB: 1Z7N). Eight histidine binding sites were identified from *T. maritima* ATP-PRTase complex. However, these were not at the histidine binding site which would be expected considering the relationship to the aminoacyl-tRNA synthetase-like and its domains function. The identified histidine binding sites were not conserved for the *L. lactis* ATP-PRTase enzyme. Here, through single point mutations, the aminoacyl-tRNA synthetase active site was identified to be responsible for

the regulatory response.[55] Overall the mode of inhibition is less well understood for this group of enzymes.

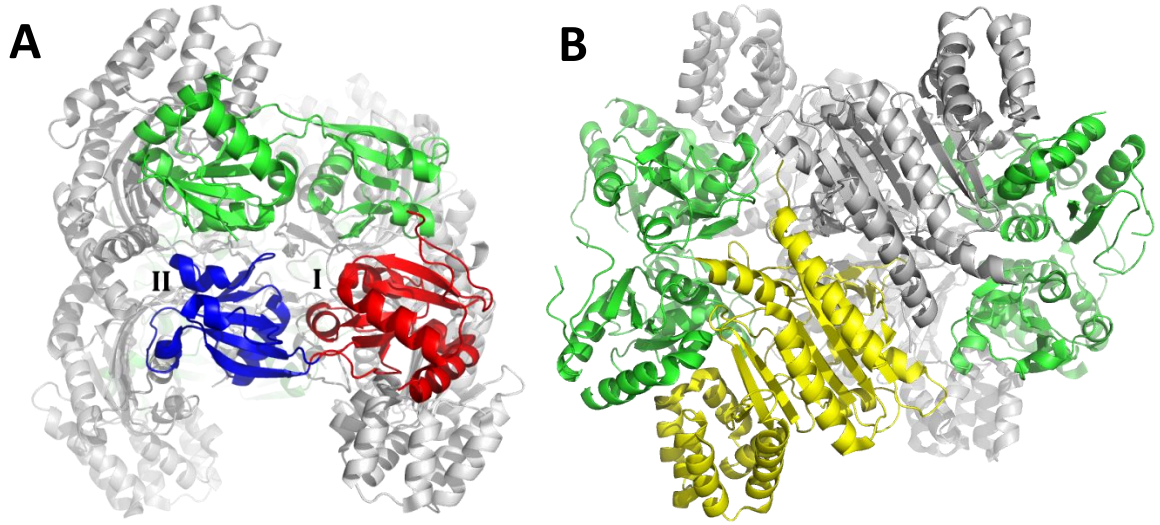


Figure 1.11. The quaternary structure of ATP-PRTase short form enzyme, from L. lactis (1H7N). ***A*** *The catalytic domain of ATP-PRTase short form comes together as a dimer, domain I (red) and domain II (blue) of one monomer are indicated and the other dimer is represented in green.* ***B*** *The aminoacyl-tRNA synthetase-like regulatory domain (single chain in yellow, remaining chains are in grey) come together in an X shaped complex with two catalytic dimers (green) at each end of the complex.*

1.9 The active site of ATP-PRTase

ATP-PRTases have a large solvent exposed active site found in the cleft between domain I and II, and the active site encompasses residues from both domains. The active site can be divided into two sites based on where these ligands bind. Residues expected to make up the ATP binding site are mostly found on domain I whereas residues involved in PRPP binding are primarily found on domain II.

A crystal structure with PRPP bound at the active site of all four chains of *L. lactis* ATP-PRTase has been reported (PDB: 1H7N).[55] In the crystal conditions ATP and *N*-1-methyl-ATP were also present; however, these ligands were not observed crystallographically. Like for Type I PRTase enzymes the 5-phosphate of PRPP is accommodated by the phosphate binding loop (Figure 1.12, **A**). This signature fold (β9–α7) varies slightly from Type I PRTase enzymes and consists of an acidic residue of either aspartic acid or glutamic acid followed by two hydrophobic residues with the remaining residues including at least one glycine and one threonine. The ribose 2- and 3-hydroxyl groups interact with Asp155 and Glu142.

Asp155 is found at the start of the phosphate binding loop and Glu142 is functionally conserved for the reported structures of both the long and short form enzymes.[47,49,54,55] Variations are observed around the pyrophosphate part of PRPP across the four active sites, potentially as a result of the lower resolution of 3.0 Å. Residue Ser140, which is highly conserved, and Lys8' from the adjacent monomer, are expected to play an important role in binding this part of the ligand.[55] Additionally, the catalytically essential Mg^{2+} cation is not observed.

The crystal structure with the PRATP product bound to the active site of *E. coli* ATP-PRTase is reported (PDB: 1Q1K), however only the PRA part of the full length structure is observed crystallographically and the remaining part of the molecule was modelled into the active site (Figure 1.12, **B**).[49] PRA part binds at the active site with the 5'-phosphate interacting with the phosphate binding loop. The ribose hydroxyl groups interact with the two conserved acidic residues Asp169 and Glu156, of which Asp169 is found at the start of the phosphate binding loop. Overall the binding mode is similar to what is observed for PRPP for this part of the molecule. The adenine ring forms hydrophobic interactions with Leu170 and Cys104 on one side and a π-π sandwich stacking interaction with the guanidine group from Arg16 on the other side. Interestingly Arg16 is not conserved across species; for example in *M. tuberculosis* ATP-PRTase an alanine is found at this position.

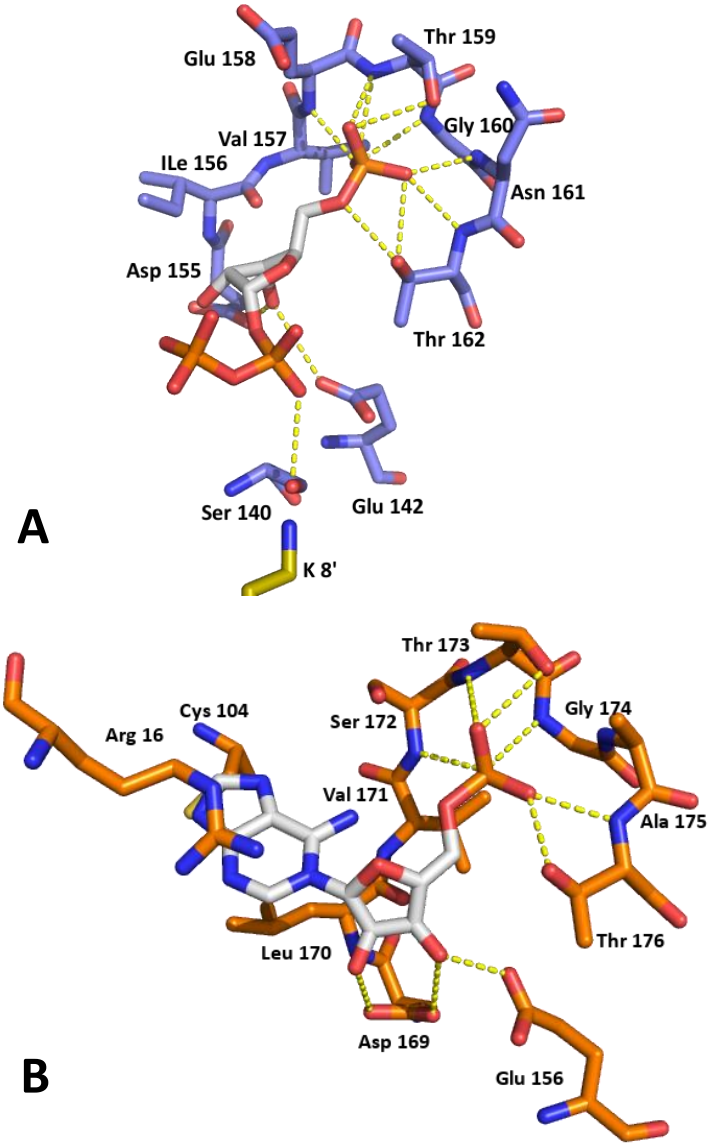


Figure 1.12. ATP-PRTase active site. ***A*** *crystal structure of L. lactis ATP-PRTase with PRPP bound. The main chain residues are in blue and residues for an adjacent monomer are in pale yellow (1H7N),* ***B*** *E. coli ATP-PRTase in brown with PRA (white) bound at the active site (1Q1K). In both structures oxygens are in red, nitrogen in blue and phosphate in orange.*

1.10 The reaction mechanism

The reaction of ATP-PRTase follows an ordered sequential reaction mechanism with ATP binding first followed by PRPP.[56] After product formation PP_i is released first, followed by the product PRATP with overall inversion of the stereochemistry at the site of the reaction.[56] The order of binding brings both substrates together before catalysis, resulting in a possible reaction mechanism being A_ND_N or S_N2-like, where nucleophilic attack and PP_i departure take place in the same reaction step. However, a general reaction mechanism has been proposed for Type I PRTases, which share large structural similarities with Type IV enzymes,

involving product formation through ribose migration.[38] This mechanism was based on crystal structures of nine different PRTases involved in nucleotide synthesis.[38,57] Crystal structures provide information on states relevant to catalysis, in a way providing snapshots along the reaction pathway. Smith *et al*. found the position of the ribose ring varied whether C1 is bonded to PP_i or the nucleophile, with little change in position of the nitrogenous base or PP_i observed in the different structures. From this, movement of the ribose ring as a carbocation intermediate was proposed, consistent with a D_N+A_N or S_N1 reaction mechanism.[38,58]

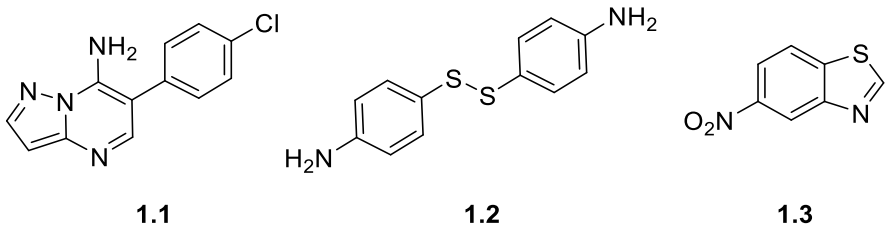
A powerful approach to characterise reaction mechanisms in enzyme catalysed reaction is through kinetic isotope effect (KIE) studies.[59] One of the challenges however, can be that events other then the chemical step of interest along the reaction pathway can mask the observation of an isotope effect. KIE studies on a number of PRTase enzymes from various host organisms including ATP-PRTase from *Salmonella typhimurium* have been reported.[58] In the case of ATP-PRTase suppression of KIEs was observed in both the forward and reverse directions, leading to inconclusive results.

KIE studies and transition state analysis on orotate phosphoribosyltransferases (OPRTases), a Type I PRTase enzyme, from a number of host organisms have been investigated.[60–62] These studies showed that the reaction catalysed by OPRTase passes through a dissociative reaction mechanism. OPRTase belongs to the Type I PRTase subfamily and supports the proposed reaction mechanism based on crystal structures for this group. Given the level of conservation around the PRPP binding site in this group of enzymes and ATP-PRTase, these findings support a dissociative reaction mechanism as a likely mechanism for the ATP-PRTase enzyme. The lack of the flexible loop and small changes around the two acidic residues at the PRPP binding site, in addition with the observation that ATP binds in the active site first, means that a direct role of the ATP participating as a nucleophile at the transition state cannot be ruled out.

1.11 Known non-natural active site inhibitors of ATP-PRTase

To date a number of non-natural inhibitors for the ATP-PRTase enzyme have been identified. A study by Dall-Larsen *et al*. has shown that pentachlorophenol, dicoumarol, and 2,4-dinitrophenol have a moderately inhibitory effect with competitive inhibition constants

(K_i) values ranging 50–400 µM for the *E. coli* enzyme.[63] Gohda *et al.* performed a three dimensional pharmacophore search technique to find inhibitors based on the product of the reaction (PRATP) for ATP-PRTase from *Arabidopsis thaliana*.[64] This led to the discovery of biaryl containing inhibitors, with the most potent inhibitors containing a disulphide-bond between the two rings (Figure 1.13, **1.1** and **1.2**). Inhibition by this class of compounds can be explained through covalent binding to cysteine residues present in the enzyme, making them non-specific and for that reason these compounds were not ideal for further drug development. In 2008 Cho *et al.* used a virtual screening approach to try and find inhibitors for *M. tuberculosis* ATP-PRTase.[65] In this study, over 500,000 lead and drug-like compounds were docked into the active site of *M. tuberculosis* ATP-PRTase, which led to the discovery of several 1–10 µM inhibitors in the initial screen. One of the strongest inhibitors contained a nitrobenzothiazole core (Figure 1.13, **1.3**) which has similarities to the dinitrophenol compound found for the *E. coli* enzyme. Additional structure activity relationship (SAR) studies by virtual screening were used to find additional sets of compounds based around this nitrobenzothiazole core group, with two compounds demonstrating inhibition of bacterial growth in a whole cell assay on *Mycobacterium smegmatis*.



1.1 **1.2** **1.3**

*Figure 1.13. Examples of non-natural inhibitors for ATP-PRTase enzymes. Compounds **1.1** and **1.2** exemplify biaryl containing compounds.* [64] ***1.3** the nitrothiazole core found in a virtual screening afford by Cho et al.* [65]

1.12 Outline of thesis

ATP-PRTase is responsible for the first step in the biosynthesis of histidine, where it controls the flux into the pathway. The absence of this enzyme in humans and the essential nature of this pathway to pathogenic organism makes this enzyme an interesting drug target. The kinetic properties of the two distinct forms of the ATP-PRTase enzyme have been investigated on the substrate level; however, the current knowledge regarding substrate binding and details of the species involved during catalysis are limited.

This project aims to gain a better understanding of substrate binding at the active site of ATP-PRTase, the reaction mechanism and transition state of ATP-PRTase enzymes, and the inhibition of ATP-PRTase through non-natural inhibitors.

Chapter 2 reports the crystallisation of *C. jejuni* ATP-PRTase in the active ATP-bound form, and the inactive closed form, in complex with histidine and AMP. A crystal structure of ATP bound to the active site of *M. tuberculosis* ATP-PRTase was obtained and the differences in binding modes are discussed.

Chapter 3 encompasses the work towards measuring kinetic isotope effects for ATP-PRTase to examine the reaction mechanism. This work includes the chemical synthesis of ^{15}N labelled adenine, coupled enzymatic synthesis of ^{15}N labelled ATP, coupled enzymatic synthesis of unlabelled and ^{3}H, ^{14}C and ^{15}N labelled PRATP, the investigation of an alternative PP_i substrate for ATP-PRTase enzymes, conditions for kinetic isotope effect measurement, and the measurement of kinetic isotope effects for *M. tuberculosis*, *C. jejuni* and *L. lactis* ATP-PRTase enzymes.

Chapter 4 discusses the generation of ground state and transition state models through density functional theory (DFT) calculation with the experimental KIEs as boundary conditions to the generated structure. A potential reaction mechanism based on the modelled transition state has also been proposed.

Chapter 5 presents the design, synthesis and characterisation of potential transition state analogues. The inhibitory properties of the designed ligands will also be discussed.

Chapter 6 covers a fragment based lead discovery effort in an attempt to find non-natural inhibitors for the ATP-PRTase enzyme from *M. tuberculosis*.

Chapter 7 provides the overall conclusion and future direction of the research.

This work was carried out on three different ATP-PRTase enzymes: two of these were from pathogenic host organisms, namely *M. tuberculosis* and *C. jejuni*, which both express the long form ATP-PRTase enzyme. ATP-PRTase from *L. lactis* was also included to investigate how the presence of an alternative regulatory domain may affect the reaction mechanism.

The *L. lactis* ATP-PRTase enzyme was chosen as it has been studied in some detail. The inhibition studies focussed on the ATP-PRTases from *M. tuberculosis* and *C. jejuni*.

Chapter 2

Structural analysis of adenosine triphosphate phosphoribosyltransferase

2.1 Introduction

This chapter discusses the structural basis for regulation of the *Campylobacter jejuni* ATP-PRTase enzyme. Two crystal structures, one with the substrate ATP bound to the enzyme, and the other with histidine and AMP, were obtained as part of this study. Prior to commencing the studies reported in this chapter, the quaternary structure of *C. jejuni* ATP-PRTase in solution was examined by Dr G. H. Mittlestädt.[66] The enzyme was studied by analytical ultracentrifugation (AUC), as well as analytical size exclusion (SEC) and SEC-coupled static light scattering techniques. This confirmed *C. jejuni* ATP-PRTase to mainly exists as a hexamer even at low enzyme concentration.

Furthermore, a crystal structure with ATP-bound at the active site of *M. tuberculo*sis ATP-PRTase was also obtained, enabling comparison of the binding mode of ATP to the *C. jejuni* ATP-PRTase enzyme.

2.2 Protein crystallisation and structure determination of *C. jejuni* ATP-PRTase

2.2.1 Crystal conditions

Suitable crystal conditions for the co-crystallisation of ATP, and histidine and AMP were identified from Clear Strategy I and II screens (Section 8.2.4.4) using a Mosquito® crystal robot. Suitable conditions were further optimised using the hanging drop vapour diffusion method to give single crystals of uniform shape and larger size. In this it was important to ensure freshly purified *C. jejuni* ATP-PRTase enzyme was used, stored no longer then 1-2 days at 4 °C, to grow single uniform crystals. Typically crystals were fully formed after 1-3 days and generally grew as cuboid-shaped crystal (Figure 2.1). Cryo-protection required particular care and the best method was found by using a 40 % glycerol, 60 % reservoir solution, slowly added directly to the crystal drop before looping. Datasets were collected with ligand density observed for ATP, and histidine and AMP (Table 2.1). This is the first reported ATP-PRTase enzyme with full ligand density observed for the ATP ligand. Data collection was carried out on the MX beamlines at the Australian Synchrotron.

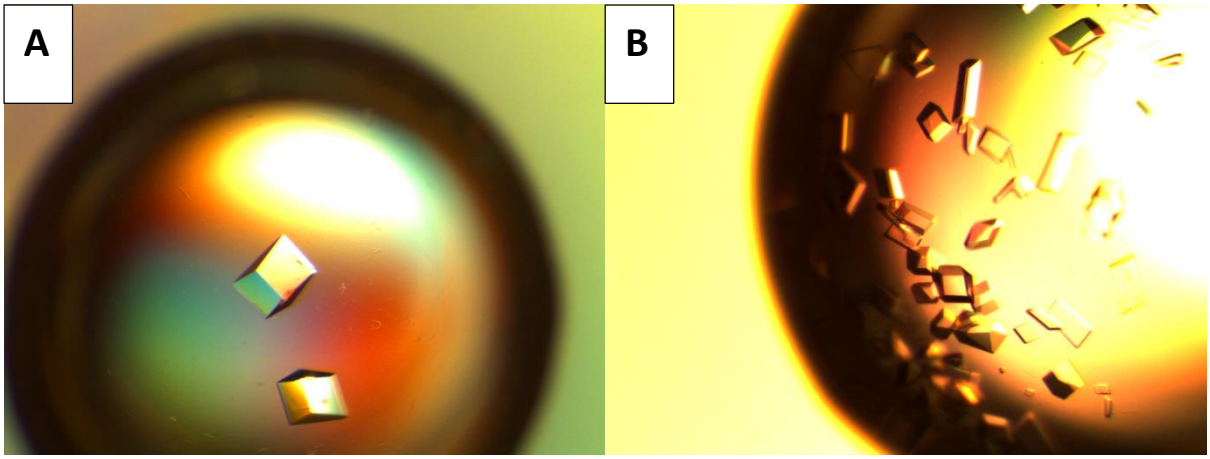


Figure 2.1. C. jejuni ATP-PRTase protein crystals. ***A*** *Co-crystallisation of ATP was achieved in 0.1 M sodium acetate, 0.1 M $MgCl_2$, 13–15 % w/v PEG 4000, at pH 5.5 and 10 mM ATP.* ***B*** *Co-crystallisation of the synergistically inhibited structure was achieved in 0.1 M Tris, 0.1 M $MgCl_2$, 13–15 % w/v PEG 4000, pH 7.5 with 1 mM histidine and 5 mM AMP.*

2.2.2 Molecular replacement

The crystal structure with the allosteric ligand histidine bound, had been previously solved for *C. jejuni* ATP-PRTase and was used in the molecular replacement strategy for the ATP, and the histidine and AMP structure.[67] A molecular replacement strategy using the full hexameric arrangement from the histidine bound structure provided a good model for the synergistically inhibited crystal structure. This initial solution was further refined with clear density for both histidine and AMP.

Molecular replacement of the ATP-bound crystal structure proved more challenging. No solution was found when the monomeric unit of the histidine bound structure was used. The final molecular replacement strategy involved several steps, initial searches involved only the catalytic core (domain I and II) to find a suitable dimeric arrangement. This structure was then used in a subsequent search to find two hexameric assemblies. This was consistent with the expected number of chains based on the unit cell dimensions. Domain III was then built back into the chains through several rounds of refinement.

In both structures ligands bound to the enzymes were added last in the refinement process to avoid introduction of model bias. Information on data collection and refinement parameters can be found in Table 2.1.

Table 2.1. Crystal parameters, data collection and refinement statistics for C. jejuni ATP-PRTase

	ATP	Histidine and AMP
Data collection		
Crystal system, space group	Triclinic, *P* 1	Monoclinic, *P* 1 2_1 1
Unit cell parameters		
a, b, c (Å)	91.67, 91.83, 154.90	91.87, 124.91, 92.81
α, β, γ (°)	101.11, 95.21, 118.14	90.00, 115.86, 90.00
Resolution range (Å)	50.00-2.20 (2.24-2.20)	50.00-1.98 (2.01-1.98)
Measurements	839825	493057
Unique reflections	212958	127994
Redundancy	3.9 (4.0)	3.9 (3.9)
Completeness (%)	98.1 (97.3)	98.1 (96.9)
$I/\sigma(I)$	9.8 (1.6)	14.6 (2.9)
R_{mean}	0.108 (1.189)	0.090 (0.623)
$CC_{1/2}$	0.982 (0.568)	0.998 (0.837)
Wilson *B*-value (Å^2)	36.8	21.3
Mathews coefficient	2.72	2.37
Refinement		
Resolution (Å)	48.06-2.20	41.34-1.98
R_{cryst}	0.236	0.203
R_{free}	0.259	0.220
Chain length	300	300
Observed number of residues	2x292, 2x293, 3x294, 1x295, 4x296	1x289, 1x290, 1x293, 2x295, 1x296
Water molecules	365	714
Other	12 ATP, 12 Mg^{2+}, 6 PO_4^{3-} , 1 acetate	6 His, 6 AMP, 9 Mg^{2+}, 6 PEG
Mean *B* (Å^2)		
Protein	47.13	24.58
Water	35.20	26.94
Other	55.83	45.09
Ligand	51.27	13.06, 24.81
RMSD from target values		
Bond lengths (Å)	0.009	0.009
Bond angles (°)	1.470	1.372
Dihedral angles (°)	0.073	0.077
Ramachandran		
Preferred (%)	98.22	98.36
Allowed (%)	1.78	1.64
Outliers (%)	0.00	0.00
PDB entry	4YB7	4YB6

2.3 The structure of *Campylobacter jejuni* ATP-PRTase

The ATP-bound *C. jejuni* ATP-PRTase crystallised in space group P1 and the unit cell contains a total of twelve chains forming two complete hexamers (Root mean square deviation (RMSD) between hexamers 0.50 Å). A single Mg^{2+}-bound ATP molecule is found in the active

site of each chain. The second crystal structure with AMP and histidine bound at the active site and allosteric site was solved in space group $P12_11$ and contains six chains, arranged in a single hexamer. The structure features six well-defined histidine molecules bound to the regulatory domain of the hexamer and one AMP molecule per active site. There were 4-11 residues per chain that were not resolved.These residues are located at the N-terminus, as well as external loops (α1-β2, β6-α4, and β13-β14 loops).

Each individual *C. jejuni* ATP-PRTase chain is composed of 10 α helices, 16 β strands and the connecting loops (Figure 2.2). Each chain is arranged in three domains, domains I, II, and III, adopting the same overall architecture as the *M. tuberculosis* ATP-PRTase and *E. coli* ATP-PRTase.[47,49] *C. jejuni* ATP-PRTase domain I is composed of residues 0-103 and 191-225, which form a Rossmann fold, commonly found for nucleotide binding proteins.[68] Residues 104–190 form domain II, which is inserted into domain I between the two β-strands β5 and β11. Together the two domains form a bilobal core unit that contains the catalytic site in the cleft between them, where the binding of ATP and AMP is observed. The last 74 C-terminal residues form domain III. In the structure with AMP and histidine, the allosteric inhibitor histidine is bound at the interface of domain III of two different chains and is referred to as the regulatory domain.

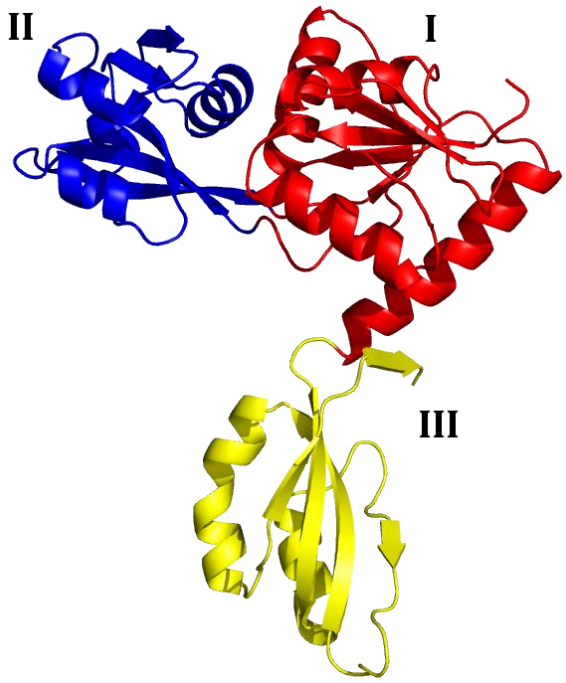


Figure 2.2. The C. jejuni ATP-PRase chain structure. Architecture of a single chain of the histidine and AMP bound C. jejuni ATP-PRTase structure is given in a three-dimensional cartoon representation. The colouring is according the three domains, domain I in red, domain II in blue and domain III in yellow.

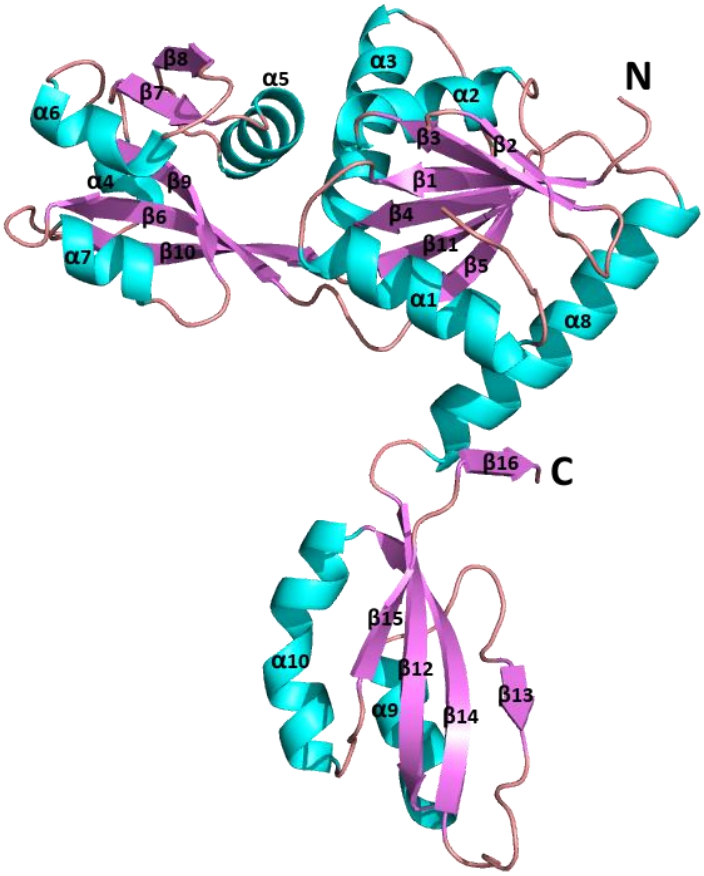


Figure 2.2 (continued) The same single chain with secondary structure elements labelled according α-helices in cyan, β-sheets in violet, and flexible loops in salmon.

The individual chains of both *C. jejuni* ATP-PRTase structures form homo-hexamers via interaction at two main interfaces. The first of these, the "dimer interface", is created by the catalytic domains of two chains with an approximate 180˚ rotational symmetry (Figure 2.3). The interface is mainly formed by interactions from β2 and β3 and the end of the α1-β2 loop on domain I to the α-helices α6 and α7 and the loops at either end of α6 (β8-α6 and α6-β9) on domain II of the adjacent monomer. In addition a small number of interactions are made between α2 to the β8-α6 loop as well. The two fold symmetry means that these interactions are repeated at the other end of the monomer. The dimer interface buries an area of 950 ± 21 Å^2 and 990 ± 77 Å^2 of each subunit in histidine and AMP, and ATP-bound forms, respectively.

The second interface, the "trimer interface" is formed around the three fold symmetry axis and pays large contributions to the overall hexameric structure. The interface is formed by interactions between the regulatory domain from different dimers: the regulatory domains

from three dimeric units come together to form two trimeric arrangements at either end thereby forming a hexamer, placing all active sites on the inside of the hollow-centred complex. Accordingly the trimer interface occurs six times per hexamer and is generated between the two different faces of neighbouring regulatory domains, so that each chain contributes to two trimer interfaces. The main point of interaction is the C-terminal β-strand β16 (residues K296-L298), which is donated into the β-sheet of the adjacent regulatory domain. The buried surface area of the trimer interface ranges from 822 ± 70 Å^2 (ATP-bound, PDB: 4YB7) to 1224 ± 21 Å^2 (Histidine and AMP, PDB: 4YB6).

The PISA complexation significance score (CSS) was calculated for both structures. The PISA CSS allows for comparison of the contribution of the interfaces relative to each other.[69] The score is defined as a maximal fraction of the total free energy of binding that belongs to the interface in a stable assembly. The dimer interface gave a CSS of 0.33 for both structures where for the trimer interface a CSS of 0.53 (ATP bound) and 0.66 (histidine and AMP bound) was found. Overall the trimer interface was classified as more significant than the dimer interface in both structures which is consistent with the observation of the hexameric assembly in solution.[66]

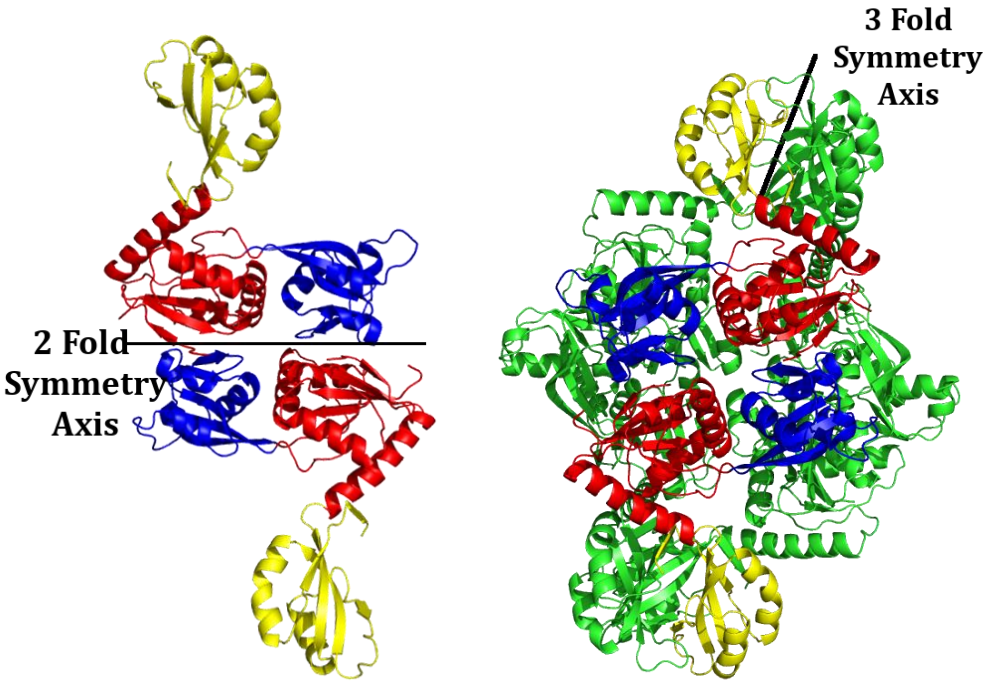


Figure 2.3. C. jejuni ATP-PRTase comes together in a hexameric arrangement with a two fold symmetry between two catalytic cores (domain I (red) and II (blue)). The trimer interface forms the three fold symmetry Axis around the regulatory domain (yellow) making the overall structure hexameric. The structure shown is from the synergistically inhibited enzyme (4YB6).

2.4 ATP binding is consistent with a sequential mechanism

In the ATP bound structure, ATP was observed for the first time in any ATP-PRTase structure to be present at full occupancy and was present in all 12 chains. In the observed binding mode ATP occupies the majority of the active site, but does not involve the conserved 14 amino acid PRTase signature motif found on domain II.[38] This fold is characteristic to type I and IV PRTase enzymes and includes the PRPP loop that is known to specifically bind phosphate-groups, or chemically similar compounds, and therefore plays a key role in substrate binding. This finding is in line with a sequential binding mechanism, in which ATP can bind first, leaving the active site open for interaction with PRPP.

ATP is accommodated by a groove, part of domain I, involving interactions with a number of residues that are in part conserved. The triphosphate moiety of ATP interacts with the side chains of Arg54 and Arg16, which form H-bonds with the γ and α-phosphate respectively (Figure 2.4). Additionally the Gln12 side chain is placed in an optimal position to interact with several of the triphosphate oxygen atoms. A determining feature of the ATP binding site is the presence of a Mg^{2+} ion in complex with the bound ATP. It is coordinated in an octahedral arrangement created by the side chains of residues Asp55 and Asp56, the β and γ-phosphate of ATP, and two water molecules. This coordination embeds the Mg^{2+} into the binding site of *C. jejuni* ATP-PRTase, providing a strong interaction with the ATP triphosphate.

The ATP ribose ring is positioned in a small cavity created by the backbone and short sidechains of Leu17, Gly73, Gly102, Ser191, Arg192 and Ala193 and the 2’ and 3’ hydroxyls form H-bonding interactions with the functionally conserved residue Asn75. The adenine is held in place by hydrophobic interactions with the side chain of Arg16 on one side and Leu170 and Cys104 on the other side. Interaction of Arg16 is through a π-π sandwich stacking interaction with the adenine base. The hydrophobic pocket allows the adenine to adopt two different orientations. Adenine was observed in either a catalytic relevant position, placing the N1 nitrogen close to the expected PRPP binding site, or flipped by an 180˚ rotation, moving N1 away from the possible reaction site.

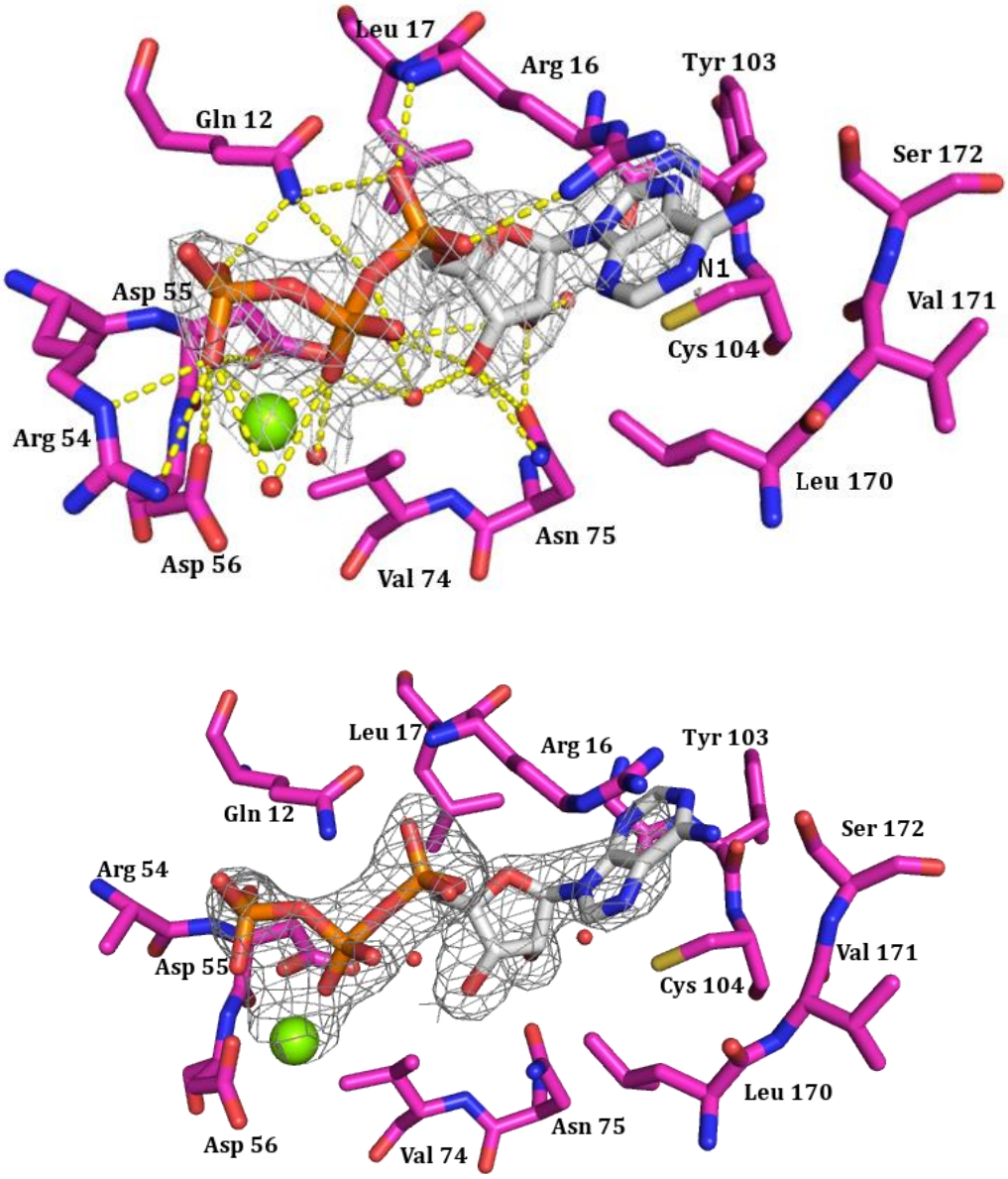


Figure 2.4. The ATP binding site. ATP was observed to adopt two binding modes in structure 4YB7. Stick representation of the ATP molecule (white) surrounded by the active site residues (pink). Water molecules and the Mg^{2+} ion are represented as spheres. The Omit map (Fo-Fc) is displayed in grey. Mg^{2+} is shown as a sphere (lime). Heteroatoms are coloured according to element: oxygen (red), nitrogen (blue), phosphorous (orange).

2.5 AMP-binding mode confirms competitive characteristic towards ATP

AMP was found to interact almost entirely with residues of the PRPP-binding loop, Asp169-Thr176 (Figure 2.5). Residues Ser172–Thr176 tightly coordinate the phosphate group of AMP by H-bonding interactions with their backbone nitrogen atoms and the side chain oxygen atoms of Ser173 and Thr176. H-bonding interactions also occur between the ribose hydroxyl groups and the side chain of the conserved residue Asp169. The AMP adenine is

positioned over Leu170, as described for ATP, but with an inverted orientation. A close to identical binding mode was reported for the AMP-bound crystal structure of *E. coli* ATP-PRTase (Section 1.8.1.2).[49] Evidently AMP-binding occurs in both the ATP and the PRPP-binding regions of the active site, which strongly supports the competitive inhibition of AMP towards both substrates observed in the kinetic analysis (Figure 2.6).

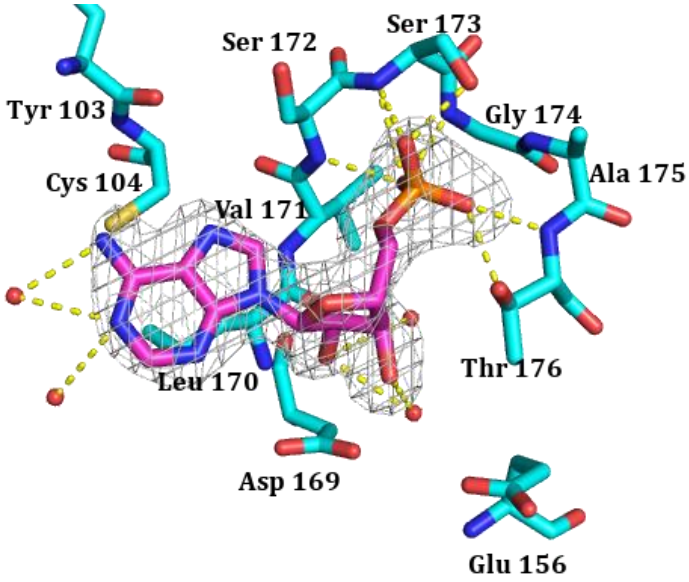


Figure 2.5. AMP binding at the active site. AMP (pink) as presented in crystal structures 4YB6 is displayed as sticks (cyan) with the omit map (Fo-Fc) displayed in grey. Mg^{2+} is shown as a sphere (lime). Heteroatoms are coloured according to element: oxygen (red), nitrogen (blue), phosphorous (orange).

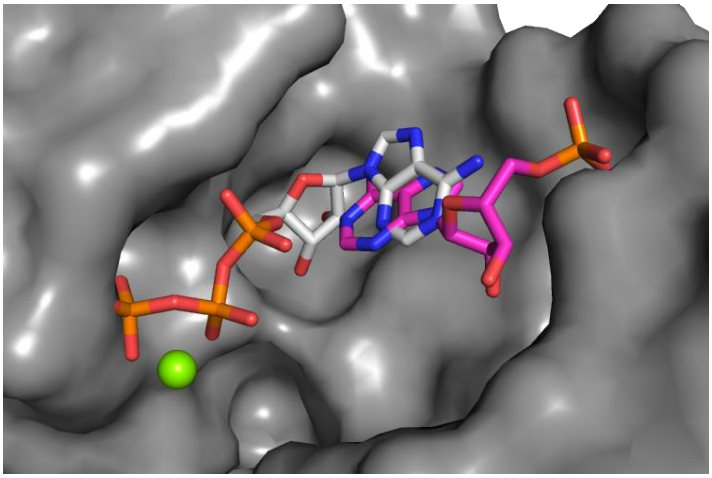


Figure 2.6. Overlap of AMP and ATP-binding at the active site. AMP (pink) and ATP (white) as presented in crystal structures 4YB6 and 4YB7 respectively are displayed as sticks in the active site of C. jejuni ATP-PRTase (grey) in surface representation. Mg^{2+} is shown as sphere (lime). Heteroatoms are coloured according to element: oxygen (red), nitrogen (blue), phosphorous (orange).

2.6 The histidine binding site of *C. jejuni* ATP-PRTase

Histidine-binding occurs at the interface between two adjacent regulatory domains in their trimeric arrangement. Consequently one histidine is bound per chain, but each chain provides two different faces for the interaction with histidine. The histidine carboxyl and amino groups are coordinated via backbone interactions with a highly conserved binding loop between α9 and β13 of one regulatory domain (consensus sequence: PGXXXPT, *C. jejuni* ATP-PRTase residues Pro246–Thr252), while the imidazole side chain is inserted into a small cavity created by the side chains on the β-sheets of the other regulatory domain, involving interactions with His232', Ser288', Leu290', and a conserved water molecule. A very similar binding mode for histidine was found in the of *M. tuberculosis* ATP-PRTase structure (PDB: 1NH8)(Section 1.8.1.1).[47]

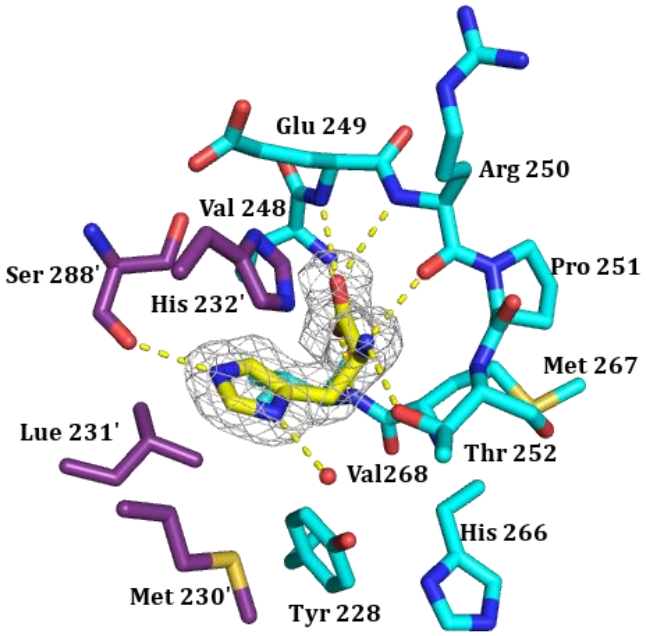


Figure 2.7. Histidine-binding site. Binding mode of the allosteric inhibitor histidine as presented in the histidine and AMP bound C. jejuni ATP-PRTase crystal structures. The omit map (Fo-Fc) is displayed in grey. Heteroatoms are coloured according to element: oxygen (red), nitrogen (blue), phosphorous (orange). The main chain is in part omitted for display purposes. Residues from the main chain are in cyan and residues from the adjacent monomer are in purple.

2.7 *C. jejuni* ATP-PRTase exists in two distinct conformational states

The ATP-bound structure (4YB7) differs considerably from the histidine bound structures (4YB6) in its overall hexamer-conformation. This is apparent from the differences in the

orientation of the individual chains towards each other and the consequential reorganisation of their corresponding interfaces.

The two hexamers found in the ATP-bound structure have an overall wider and more rounded appearance when compared to the inhibitor-bound structures (Figure 2.8. **A** and **B**). The change in diameter reflects this change which is decreased by an average of 77.4 ± 0.7 Å for the ATP-bound (Figure 2.8, **C**) structure to 71.7 ± 0.1 Å for the histidine and AMP bound structure (Figure 2.8. **D**). Overall the histidine bound structures appear longer and more ellipsoid. The change from the active to the inactive conformation is visible by the rotational twist of the regulatory domain relative to the catalytic core by approximately 30˚. This change is driven by the binding of histidine at the interface of neighbouring regulatory domains, which pulls the monomers closer together, significantly increasing the buried surface area of the trimer interface from 822 ± 70 (ATP-bound, PDB: 4YB7) to 1224 ± 21 $Å^2$ (Histidine and AMP, PDB: 4YB6).

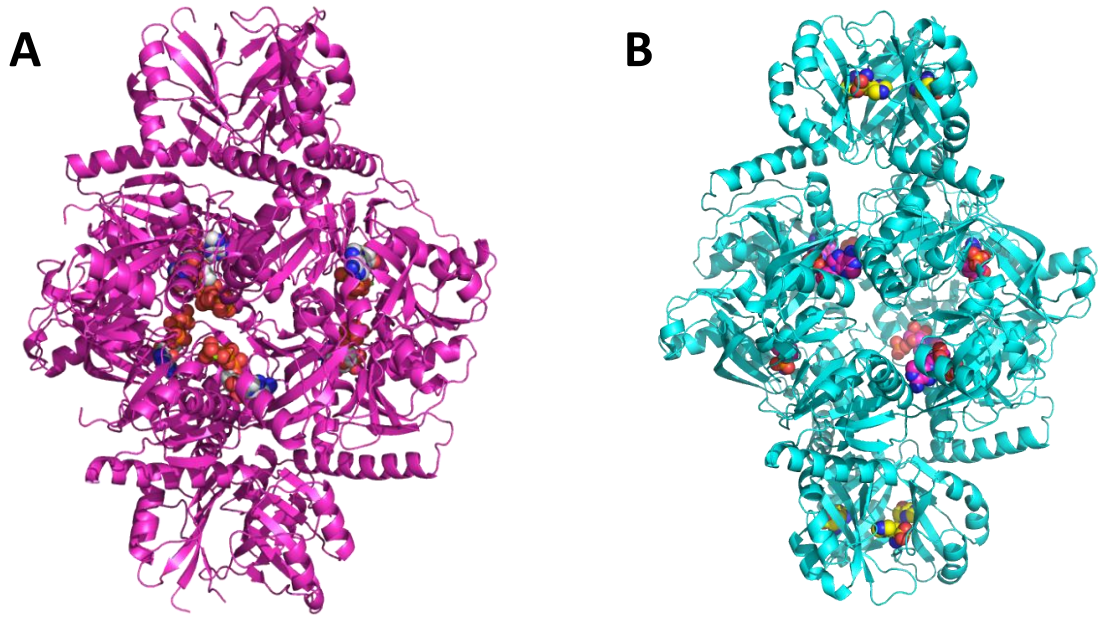


*Figure 2.8. Comparison of C. jejuni ATP-PRTase hexamer conformations. **A** shows the open ATP (white) bound structure(chains in purple) and **B** shows the histidine (yellow) and AMP (pink) inhibited structure(chains in cyan). Histidine binding induces a conformational change that results in a more narrow structure making it appear longer and ellipsoid.*

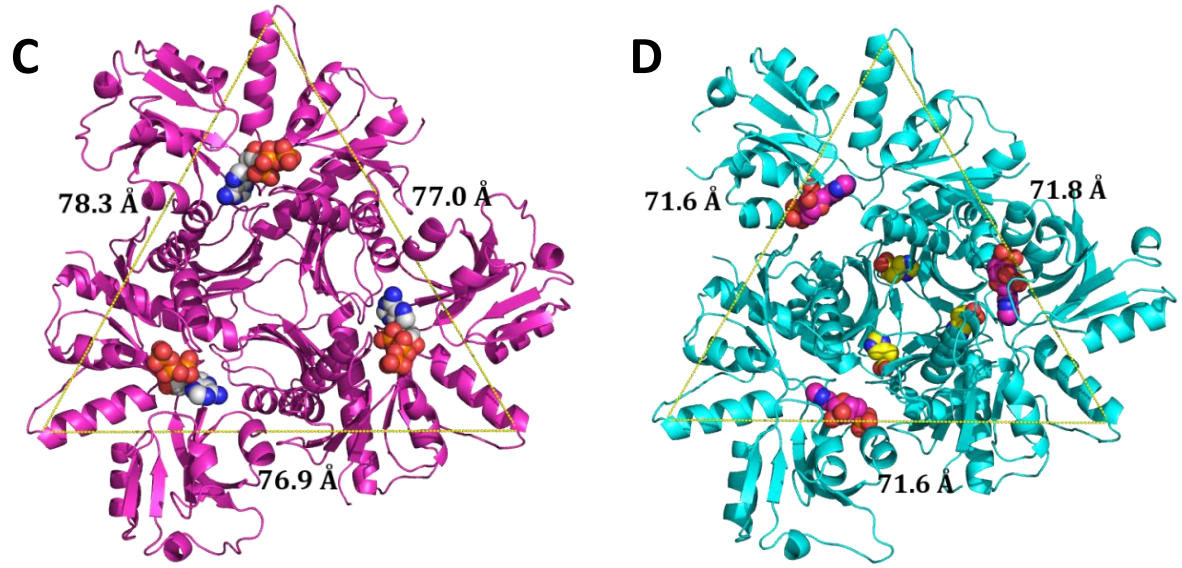


Figure 2.8. (continued) ***C*** *and* ***D*** *are cross sections looking down the trimer interface. The conformational change going from an open more round structure (**C**) to a more narrow inhibited structure (**D**). The distance, yellow dashed line, between active sites is shown to further visualise this change.*

Comparison of the occupied allosteric site, with histidine bound, to the unoccupied allosteric site of the ATP-bound structure reveals clear differences between the two conformations. This shows in the distance and orientation of the two faces of the binding site relative to each other. Although the conserved histidine binding loop does not change its conformation, it is significantly displaced from the adjacent monomer's binding face (Figure 2.9, **A**). Moreover in the presence of histidine a tight hydrogen-bonding network is formed between all three neighbouring regulatory domains via the residues Thr252, Tyr228 and His266, a water molecule and the bound histidine molecules, strengthening the interface (PDB: 4YB6) (Figure 2.9, **B**).

While histidine binding directly and prominently affects the trimer interface, the catalytic domains also contribute to the overall changes in the hexamer. The large conformational change of the regulatory domains does not affect the conformation of the regulatory domains (I and II) (1.02 ± 0.05 Å RMSD between residues 4–225 of AMP and histidine bound, and ATP-bound *C. jejuni* ATP-PRTase). However, changes in the dimer interface are evident when comparing the AMP and histidine bound with the ATP-bound structures (Figure 2.9, **C**). The two chains of a dimeric unit undergo a rotational flex with respect to each other. The majority of the central dimer interface contacts, involving residues Thr152–Leu163 (α6 and flanking loops) of the first chain and Glu37'–Ile40' (α1-β2 loop), Arg8' (β1), and Asp57' (α2)

of the second chain, stay intact upon histidine-binding, but close contacts between residues Thr176–Asn180 (α7) and Glu37'–Ile40' (α1-β2 loop) are lost as the α7 helices are shifted towards the outside of the hexamer. Additional interactions are formed on the opposite side of the dimer interface, as Asn148–Leu151 (β8) and Glu57'–Glu64' (α2) and the C-terminal ends of the α3 helices (Leu87 and Leu87') are now in close proximity to each other. This movement increases the distance between the two sites by approximately 3 Å and thereby alters the geometry of the dual-active site cavity.

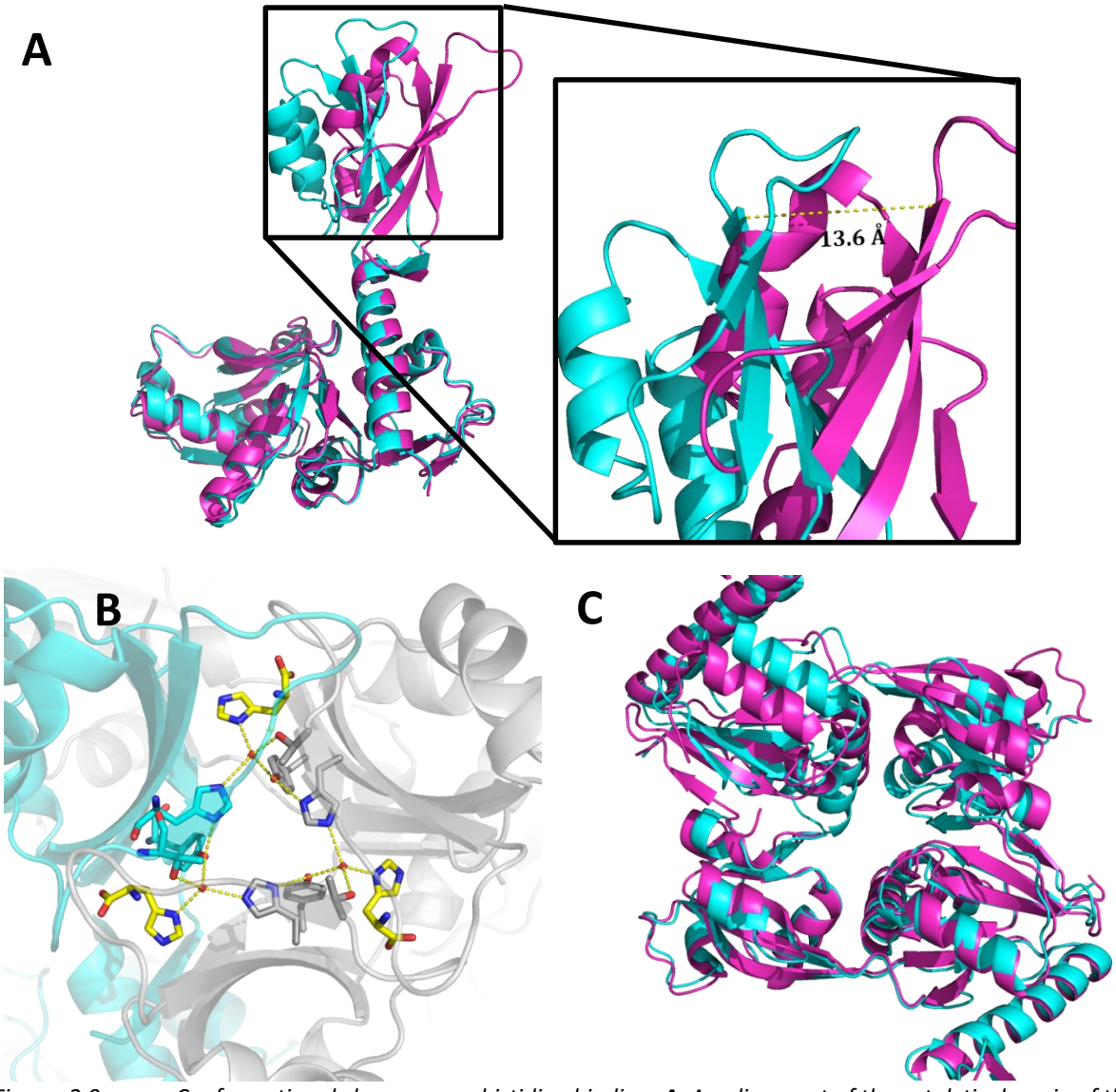


Figure 2.9. Conformational change upon histidine binding. ***A****, An alignment of the catalytic domain of the monomeric units (residues 5–225) of the active, ATP-bound form, to the inactive, histidine bound forms (4YB5 and 4YB6) reveals the large conformational change of the regulatory domain of 13.6 Å upon histidine binding.* ***B,*** *Visualisation of the formed H-bonding network associated with the binding of histidine(yellow).****C*** *Close up of the dimer interface, overlay of a single catalytic core (domain I and II) of the ATP bound structure (cyan), with the catalytic core of the AMP and histidine bound structure(purple) reveals a significant change between the two structures.*

2.8 *M. tuberculosis* ATP-PRTase structure with ATP-bound

2.8.1 Cloning, expression and purification of *M. tuberculosis* ATP-PRTase

The kinetic and structural features of *M. tuberculosis* ATP-PRTase have been studied in detail previously.[47,70] A small modification was made to the previously published sequence around the thrombin cleavage site to include a tobacco etch virus (TEV) cleavage site more commonly used within the Parker laboratory. Amplification of the gene of interest was successfully achieved from genomic DNA as determined by agarose gel electrophoresis (Figure 2.10). Both the linearized vector and PCR product were then purified by agarose gel electrophoresis.

Cloning of *M. tuberculosis* ATP-PRTase was achieved through the use of the in-fusion® cloning system outlined in Section 8.1.13. The linear product was cloned into the donor vector pET-28a. The product was transformed into Stella™ cells for DNA propagation and positive clones were determined by colony PCR and the gene encoding *M. tuberculosis* ATP-PRTase was successfully sequence verified in the donor vector.

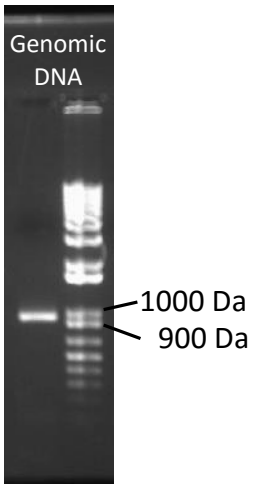


Figure 2.10. Agarose gel showing the amplified gene had the right size ~950 Da.

The plasmid was transformed into two different expression cell lines to test for soluble expression (Figure 2.11). The *M. tuberculosis* gene was transformed into *E. coli* BL21(DE3) Star cells, and *E. coli* BL21(DE3) Star pLysS cells. Both cell lines gave soluble expression and expression of the protein of interest was indicated by retention on an immobilised metal-affinity column (IMAC) and later confirmed by mass spectrometry. The *E. coli* BL21(DE3) Star pLysS cells gave better soluble expression by interpretation of SDS-page and was used in an

up-scaled growth. Further details on expression and purification of the protein can be found in Section 8.2.3.

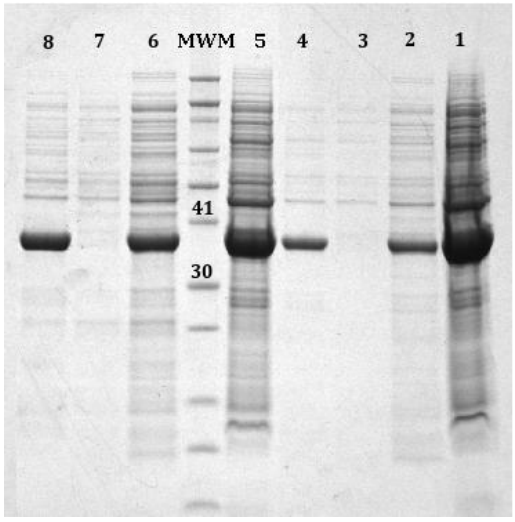


Figure 2.11. Expression test of M. tuberculosis ATP-PRTase in E. coli BL21(DE3) Star cells: ***1*** *cell pellet,* ***2*** *supernatant,* ***3*** *flow through spin Hitrap column,* ***4*** *elution spin Hitrap column and E. coli BL21(DE3) Star pLysS cells:* ***5*** *cell pellet,* ***6*** *supernatant,* ***7*** *flow through spin Hitrap column,* ***8*** *elution spin Hitrap column.*

2.8.2 *M. tuberculosis* ATP-PRTase crystal conditions

It was found that an improvement on the crystal shape and diffraction quality could be obtained by cleaving the N-terminal histidine tag. *M. tuberculosis* ATP-PRTase enzyme crystals in the apo form were grown according to published conditions using the hanging drop vapour diffusion method (Section 8.2.4.2).[47] Suitable conditions for the co-crystallisation of ATP required the addition of 10 mM ATP and a slight dilution (by 1-10%) of the condition reported for the apo form. Crystals formed after 12 to 48 hours and grew as uniform triangular prisms of various size (Figure 2.12). This resulted in clear density for ATP bound at the active site *M. tuberculosis* ATP-PRTase not previously observed. Information on the collection and refinement parameters can be found in Table 2.2.

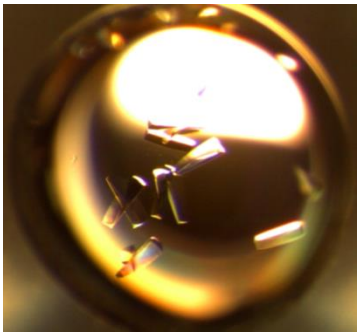


Figure 2.12. Image of M. tuberculosis ATP-PRTase crystals.

Table 2.2. Crystal parameters, data collection and refinement statistics for ATP bound M. tuberculosis ATP-PRTase crystal structure.

	ATP
Data collection	
Crystal system, space group	Trigonal, $H3_2$
Unit cell parameters	
a, b, c (Å)	134.76, 134.76, 111.00
α, β, γ (°)	90, 90, 120
Resolution range (Å)	40.99–2.40
Measurements	303991
Unique reflections	15289
Redundancy	20.0 (20.0)
Completeness (%)	100 (100)
$I/\sigma(I)$	20.0 (2.4)
R_{mean}	0.143 (1.595)
$CC_{1/2}$	0.999 (0.773)
Wilson *B*-value (Å^2)	46.5
Mathews coefficient	3.16
Refinement	
Resolution (Å)	40.99–2.40
R_{cryst}	0.222
R_{free}	0.256
Chain length	287
Observed number of residues	281
Water molecules	48
Other	9
Mean *B* (Å^2)	
Protein	53.79
Water	51.39
Other (SO_4^{3-}, Mg^{2+})	96.39
Ligand (ATP)	38.72
RMSD from target values	
Bond lengths (Å)	0.0194
Bond angles (°)	1.8505
Dihedral angles (°)	0.1067
Ramachandran	
Preferred (%)	98
Allowed (%)	1
Outliers (%)	1

2.8.3 The binding mode of ATP

Binding of ATP to the active site of *M. tuberculosis* ATP-PRTase did not significantly alter the conformation of the monomeric unit of *M. tuberculosis* ATP-PRTase in the apo form (1NH7). An overlay of the apo-crystal structure (PDB: 1NH7) and the ATP bound crystal structure gave a RMSD of 0.528 unit with only small changes observed around the active site. In addition to ATP, three sulphate ions were also observed in the PRPP binding site consistent

with the previously reported apo structure. The close resemblance in quaternary structure further supports that the ATP bound structure represent the active form of the enzyme.

ATP binding is accommodated by domain I, involving interaction with a number of residues that are in part conserved (Figure 2.13). The adenine ring is positioned over Val155, which forms the main interaction with the adenine ring and the active site. The adenine group is bound in what is considered the catalytically relevant position with the N1 position towards the PRPP binding site. The ribose ring is accommodated by a small cavity formed by Ala11, Leu12, Asp70 and Ser90, the ribose 2' and 3' hydroxyl groups form H-bonding interactions with the strictly conserved Asp70, a H-bond is also made to the side chain of Ser90. There are two Mg^{2+} ions found at the active site which both interact with the triphosphate group of ATP. One Mg^{2+} ion is positioned towards the centre of the active site with the α, β and γ phosphates coordinated around the metal ion and a further three water molecules in a octahedral configuration. The second Mg^{2+} ion is found at the dimer interface and interactions are made to the γ phosphate of ATP. This interaction is mirrored on the other side by the adjacent active site effectively sharing this Mg^{2+} ion between two ATP ligands. The γ phosphate further interacts with the conserved Arg49. Furthermore, Lys 9' from the adjacent monomer also interacts with this position.

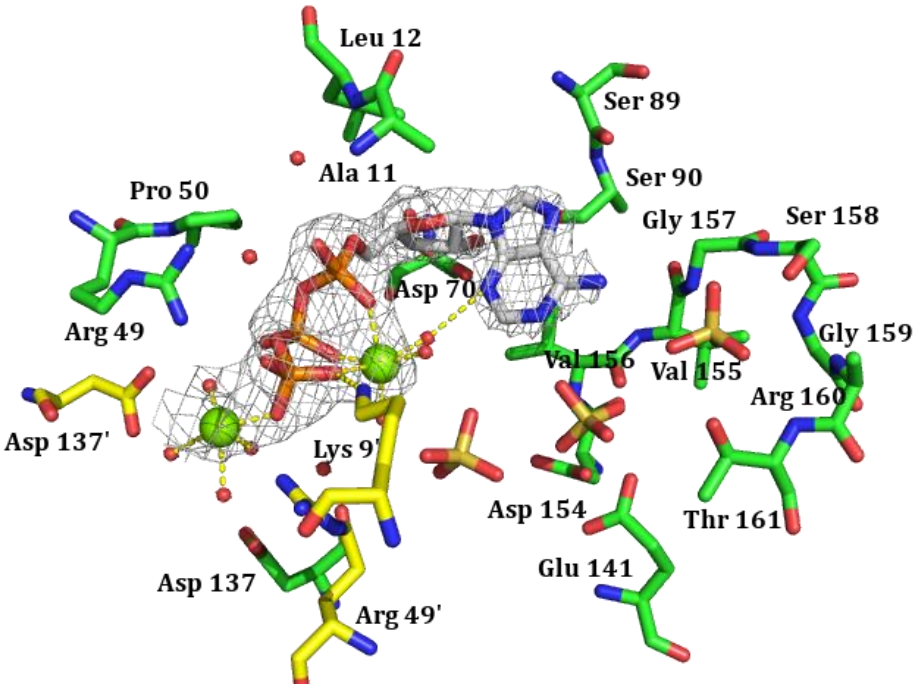


Figure 2.13. The binding mode of ATP for M. tuberculosis ATP-PRTase. Stick representation of the ATP molecule (White) surrounded by the active site residues (Cyan). Water molecules and the Mg^{2+} ion (Lime) are represented as spheres and the Omit (Fo-Fc) is displaced in grey. Heteroatoms are coloured according to element: oxygen (red), nitrogen (blue), phosphorous (orange).

Comparison of the active site of *M. tuberculosis* and *C. jejuni* ATP-PRTase displays considerable differences in amino acid composition, in particular around the ATP binding site. For example, Arg16 which was shown to be directly involved in adenine binding in the *C. jejuni* enzyme is not present in the M. tuberbulosis enzyme where an alanine (Ala12) is found. Additionally the double Asp55–Asp56 motif responsible for coordinating magnesium in *C. jejuni* ATP-PRTase is not found for the *M. tuberculosis* enzyme. Given these changes remarkable similar binding modes for ATP were found between the two structures. In particular around the adenosine part of ATP (Figure 2.14). Small differences are observed around the triphosphate part of the molecule. In *C. jejuni* ATP-PRTase the triphosphate part is held in place through coordination to a single Mg^{2+}, which in turn is held in place by the double aspartate motif, where in the *M. tuberculosis* enzyme a proline is found and overall fewer direct interactions with the active site are made.

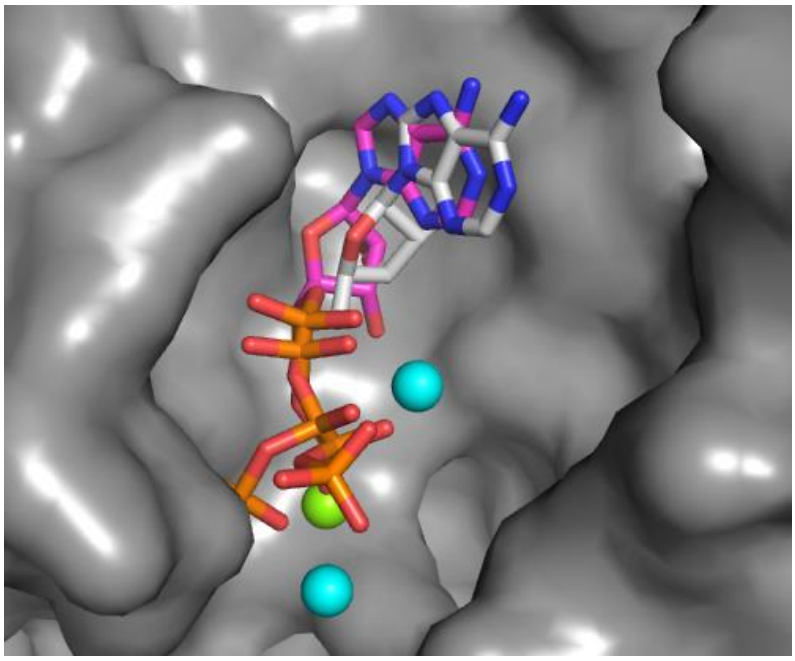


Figure 2.14. Overlay of domain I and II of ATP bound to M. tuberculosis (ATP in pink and Mg^{2+} in cyan) and C. jejuni ATP-PRTase (ATP in white and Mg^{2+} in lime) enzyme, surface is from the M. tuberculosis ATP-PRTase enzyme.

2.9 Discussion

ATP-PRTase acts as the gateway into histidine biosynthesis and as such plays an important metabolic role. *C. jejuni* ATP-PRTase shares many biochemical and structural properties with

the long form ATP-PRTase enzymes that have been studied from other sources. A competitive behaviour of AMP towards both substrates has been observed for *C. jejuni* ATP-PRTase which has also been shown for other ATP-PRTase enzymes.[56,68,71] This is consistent with the active site binding of AMP where it is observed to be present in both the ATP and PRPP binding site.

The clear observation of ATP at full occupancy for the first time in the active site gave further insight into the ATP binding site. The binding site of PRPP can be clearly inferred from this structure and from the positioning of AMP in the active site in the AMP and histidine bound structure. Additionally Mg^{2+} was observed bound to a double aspartate motif (Asp55–Asp56) underlining the importance of Mg^{2+} for the ATP-PRTase reaction. Although clearly functional, this metal-binding site along with other identified interactions is not conserved. In the active site of *M. tuberculosis* ATP-PRTase a proline (Pro48) is found in place of the double aspartate motif. Triphosphate binding is still associated with Mg^{2+} but no strong interactions are made to the active site such as what is seen in the *C. jejuni* ATP-PRTase structure.

Solution studies had shown that *C. jejuni* ATP-PRTase exists as an active hexamer in solution. Analysis of the interfaces of the crystal structures gives further support for the existence of the ATP-PRTase enzyme as a hexameric species. Two main interfaces make up the hexamer, the dimer interface and the more extensive trimer interface. Based on PISA complexation significance score the trimer interface was found to be more significant for both the open, ATP bound, and closed, histidine and AMP bound, forms of the enzyme.

In the ATP-PRTase hexamer all active sites are positioned on the inside. This poses the question of how the charged substrate and product molecules then enter or leave the enzyme with the *C. jejuni* ATP-PRTase identified as a functional hexamer. The hexameric structures of all structurally characterised ATP-PRTase long form enzymes possess the same large openings between neighbouring domains, which would likely permit access to the interior of the complex (Figure 2.15, **A**). Another possible entrance for substrate molecules is present in form of a small opening in the dimer interface, observed in the ATP-bound structure (4YB7) obtained as part of this study, which leads directly to the active site and thus may function as a substrate channel. It should also be noted that all chains observed in

the hexameric structures of *C. jejuni* ATP-PRTase differ marginally from each other, but more so in the ATP-bound crystal resulting in the loss of symmetry.

The structures of the hexameric *C. jejuni* ATP-PRTase provide insight into the changes that are associated with histidine binding and how histidine binding can contribute to the low activity. Histidine binds at the interfaces of the regulatory trimers, found at either end of the core and clearly constricts movement of this part of the protein (Figure 2.15, **B**). Upon binding, these regulatory trimers appear to rotate relative to the catalytic hexameric core and come closer together stabilising the hexamer. Some limited evidence of the stabilisation of the protein comes from thermal stability experiments performed on *C. jejuni* ATP-PRTase, which showed a significantly increased (5 ˚C) melting temperature for the inhibitor-bound enzyme.[66] The large active site of this protein is located at the hinge between domains I and II, the conformation and conformational flexibility at this position of the protein is likely to impact on catalysis and this may be restricted or altered by the structural changes associated with histidine binding. Furthermore, the rotation of the regulatory domain relative to the catalytic core reduces the size of the entrance into the hexamer described earlier. The ability of AMP to bind in presence of histidine suggests that access to the active site is not completely closed off even in the inhibited hexameric state and may account for the observed residual activity observed.

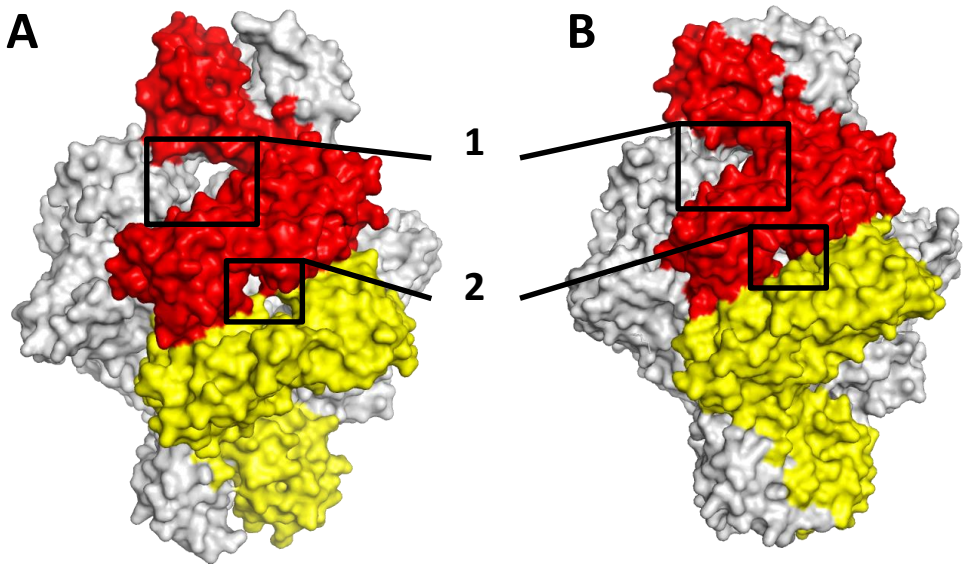


Figure 2.15. Active site access. ***A*** *A surface representation of the active ATP bound form of C. jejuni ATP-PRTase (4YB7). A large cavity can be found between the regulatory domain,* ***1****, and the catalytic core. Furthermore there is a small opening visible at the dimer interface,* ***2****, potentially providing an alternative entrance to the active site.* ***B****. A surface representation of the inhibited C. jejuni ATP-PRTase (4YB6) enzyme, access to the active site is reduced upon tightening of the regulatory domains.*

In conclusion, the ATP-PRTase enzyme from *C. jejuni* has been shown to exist as an active hexamer and a new mode of the allosteric inhibition can be proposed based on the presented crystal structures. The ATP-PRTase enzyme has been shown to exist in two states the open “active” state and the closed “inactive” state. The active state has a more open hexameric structure to allow for access of substrates and products in and out of the hexameric core. Upon histidine binding the regulatory domains come closer together with mediation of the regulatory signal through a rotational movement relative to the catalytic core. This rigidification reduces the access to the interior of the hexamer limiting substrate access. This “twisting the tail” does not result in obvious changes around the active site of ATP-PRTase, however, changes are observed around the dimer interface increasing the distance between the two catalytic sites. The lower catalytic rate may be a result of the reduced access into the interior and the reduced dynamics by restricting motion of the catalytic domains and the structure overall.

Chapter 3

Measuring kinetic isotope effects for ATP-PRTase enzymes

3.1 Kinetic isotope effects and reaction mechanisms for enzymes

3.1.1 Transition state theory

The theory of absolute reaction rates, or transition state theory (TST), was developed in 1935 simultaneously by Eyring and by Evans, and Polanyi.[72] The theory assumes that the reactants are in equilibrium with an activated complex or transition state, and the decomposition of the complex is the rate controlling step for the overall process.[72–74] Although TST is not perfect, it describes many aspects of the process well, including reaction rates, transition states (TSs), and the energy needed to achieve the TS. TST also provides a framework to explain the observation of kinetic isotope effects.[75] TST fails for some high temperatures reactions, and reactions involving small particles. Alternative theories to describe these types of reactions have been proposed.[76,77] During the work presented in this thesis quantum mechanical tunnelling through the reaction potential energy barrier was ignored because it was considered unlikely to be involved in the reaction catalysed by ATP-PRTase.[78,79] Quantum mechanical tunnelling is usually only important at very low temperatures and for reactions involving very light atoms and is therefore not required.[78,79]

3.1.2 Potential energy surface maps

A potential energy surface (PES) describes the relationship between the energy of a molecule and its geometry.[80] This is a mathematical description of the potential energy, which is a function of the bond motion (bond stretching, bond bending and bond torsion) and van der Waals interactions of the system. A potential energy surface map can be used to describe the internal energy of a molecule undergoing a reaction. In this energy landscape, the wells represent relatively low energy structures such as reactants, products, and intermediates. The peaks represent extremely high energy structures seldom achieved during a typical reaction. The lowest point between two peaks, the saddle point, represents the lowest high-energy state which must be overcome for the transformation of one structure to another. The saddle point in this case encompasses the activation energy needed for the reaction to proceed and the structure associated with that point is called the transition state. The valleys on either side represent the lowest energy path to or from the starting material and product.

3.1.3 Reaction coordinate diagrams

A reaction coordinate diagram is a cross section of a PES map along a low energy pathway, showing a single reaction coordinate (x axis) against energy (y axis). As with a PES map, depressions represent stable structures, shallow wells represent reactive intermediates, and peaks represent transition states (TS) (Figure 3.1). Any chemical structure lasting longer than a typical bond vibration (10^{-13} to 10^{-14} s) can be considered an intermediate.[80] In these diagrams, the highest energy barrier is associated with the rate-determining step of the reaction, and needs to be overcome for the reaction to proceed.

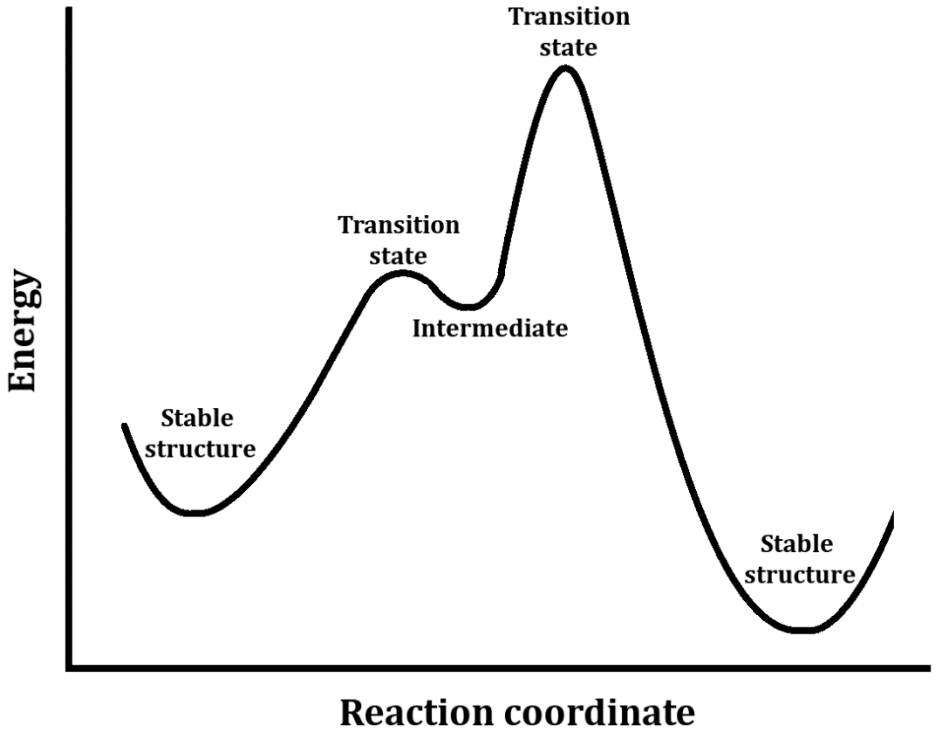


Figure 3.1. Example of a reaction coordinate diagram and the states involved to go from one stable structure to another.

3.1.4 The origin of a kinetic isotope effect

Kinetic isotope effects refer to the change in rate of a chemical reaction when one of the atoms is substituted with one of its isotopes. Isotopic molecules traverse the same potential energy surface as each other. The observation of a KIE arises from a difference in zero-point energy (ZPE) for the isotopic molecules. In the case where a bond is broken, the stretching vibration of that bond defines the reaction coordinate. A Morse potential can be used to

show the relationship between bond strength and bond length (Figure 3.2). A Morse potential diagram is a plot of the potential energy as a function of the distance between two atoms. The shape of this curve is indicative of an anharmonic oscillator.

A chemical bond vibrates in a manner constrained by the Morse potential in that a system displaced from equilibrium responds with a restoring force which is proportional to the displacement.[80] The energies of these vibrations are not continuous but are instead quantised. The energy associated with n=0 is referred to as the ZPE and is the lowest vibrational state of a molecule. At room temperature, most bonds of a molecule are in this lowest energy state. At low energy the harmonic oscillator approximation can be used to explain this vibration (Equation 1). In this equation, *h* is Planck's constant, v is the vibrational frequency of the bond, *n* is an integer representing the vibrational mode and E_n is the quantised energy of that mode.

$$E_n = \left(n + \frac{1}{2}\right) hv \qquad \text{Equation 1}$$

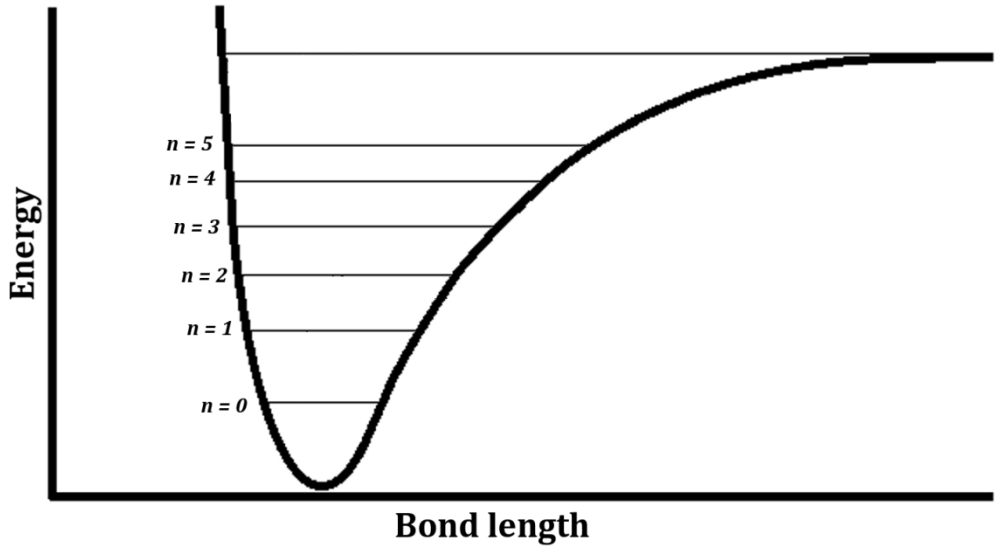


Figure 3.2. A Morse plot can be used to show the effect of different vibrational modes. n shows the quantised state of the bond.

The fundamental frequency v of a bond can then be found by Equation 2. In this equation, c is the force constant of the bond, which is the same for isotopic atoms, and μ is the reduced

mass. Reduced mass is the effective inertial mass in equations of motion for a many-particle system.[81] The value for the reduced mass itself can be found by Equation 3 where m_1 represents the weight of the anchoring atom(s) and m_2 gives the mass of the atom of interest.

This shows that the frequency is directly proportional to the square root of the force constant for the bond, and inversely proportional to the square root of the reduced mass. For instance, when a C–H bond is replaced by a C–D bond, the stretching frequency is lower because of the increased mass. The ZPE for that bond is therefore also lower. This is depicted in a potential energy well by a lower energy rung for the C–D bond (Figure 3.3). If the energy needed to break the C–H compared to the C–D bond is considered, then the difference in activation energy will be proportional to the difference in ZPE.

$$v = \frac{1}{2\pi}\sqrt{\frac{c}{\mu}} \qquad \text{Equation 2}$$

$$\mu = \frac{m_1 m_2}{m_1 + m_2} \qquad \text{Equation 3}$$

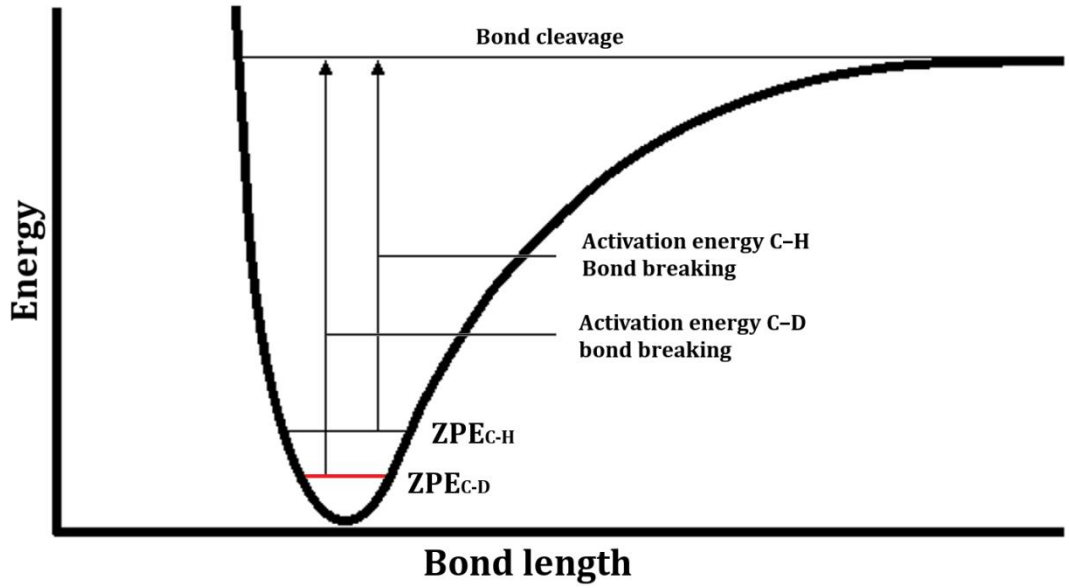


Figure 3.3. A Morse potential plot to illustrate the difference in activation energy needed for breaking a C–H versus a C–D bond. The difference in ZPE energy is shown as two rungs with the lower rung (red) for the C–D bond.[80]

To understand how this affects the rate, the Arrhenius equation is very important.[80] The Arrhenius equation for rate constants (Equation 4) allows for direct comparison of the rate of bond breaking for a C–H bond compared to that of a C–D bond. In this equation, *k* is the reaction rate, E_a is the activation energy, k_b is the Boltzmann constant, and T is the absolute temperature. Using the same example of exchanging a protium for a deuterium, k_H and k_D are the rates of reaction associated with C–H and the isotope substituted C–D. It is assumed that the pre-exponential factor A is equal for both isotopes, which is reasonable given that the isotopes follow the same reaction coordinates. The ratio of the rates of reaction then gives an approximation for the isotope effect, resulting in Equation 5.

$$k = Ae^{-\frac{E_a}{k_bT}} \qquad \text{Equation 4}$$

$$k_{light}/k_{heavy} = e^{-\frac{E_a^{light} - E_a^{heavy}}{k_bT}} \qquad \text{Equation 5}$$

With bonds featuring different isotopes, but following the same reaction path, it follows that the increase in activation energy arises only from the difference in ZPE of the chemical bond (Figure 3.3). Therefore, Equation 1 can be substituted into Equation 5, resulting in:

$$k_{light}/k_{heavy} = e^{-\frac{h(v_{Light} - v_{heavy})}{2k_bT}} \qquad \text{Equation 6}$$

Equations 2 and 3 show that the fundamental frequency of the bond is directly affected by the increase in mass of the atoms. The modified Arrhenius equation (Equation 6) thus gives the theoretical maximum isotope effect expected for bond breaking at this step.[72,74] Using the Eyring equation, in the case of breaking a C–H bond compared to a C–D bond, a theoretical k_H/k_D of ~6.7 is predicted. Isotope effects of this magnitude, however, are rarely observed. Heavy-atom isotope effects are much smaller with isotope effects <1.10 for substitutions of carbon, oxygen, nitrogen, sulphur, chloride and bromide. The difference in magnitude is directly related to the percentage mass change, as going from ^{12}C to ^{13}C is a much smaller relative weight increase than going from ^{1}H to ^{2}H.

In the C–H to C–D example it was assumed that the bond of interest would be fully broken and the observed isotope effect would arise only from the difference in ZPE of the reactant, however the effect can be more complex. Transition states are often reached with partial

bond character between multiple atoms, due to bond making and/or breaking. The bonds at these transition states, just like any other bonds, have vibrational modes and associated ZPEs. This means there is also a ZPE difference for the isotopically substituted bond. To understand how this affects the observed isotope effect, the vibrational mode undergoing a change along the reaction coordinate needs to be considered (Figure 3.4). The difference in activation energy no longer arises only from the variation in ZPE of the reactant but from the differential ZPE variation between the reactant state and the transition state:

$$\Delta ZPE_{reaction} = \Delta ZPE_{TS} - \Delta ZPE_{reactant} \qquad \text{Equation 7}$$

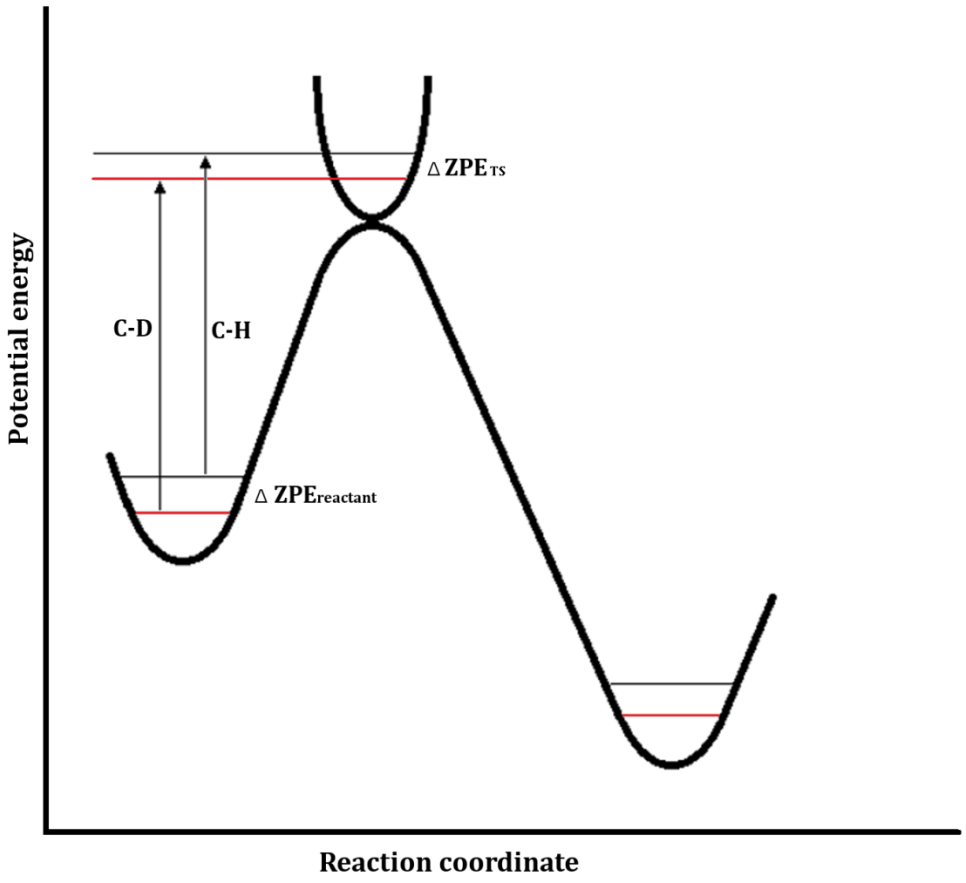


Figure 3.4. Reaction coordinate diagram *with the difference in ZPE for a C–H and C–D bond shown. Arrows indicate the activation energy needed to go from the reactant to the transition state. If a bond is not fully broken at the transition state these bonds have vibrational modes and therefore ZPE as well. The observed isotope effect arises from the differential ZPE difference. The ZPE for the C–D bond is shown in red.*

3.1.5 Secondary isotope effects

The effects discussed above have all been relating to the primary isotope effect, where an isotopically substituted atom participates directly in the bond change. When an isotopic substitution is instead made close to the atom directly involved in the reaction this gives rise to secondary isotope effects.[80] The substitution position relative to the primary position is denoted as α, β,γ, etc. Secondary isotope effects typically arise from hybridisation changes that occur during reactions, or when the electrons of a sigma bond are involved in hyperconjugation.[82] As for primary isotope effects, the change in vibrational frequencies, and therefore the change in ZPE, is responsible for the observed effect.

During a hybridisation change there are a limited number of vibrational modes that undergo a large change. Those that do undergo a large change include the stretching modes, as well as the in- and out-of-plane bending vibrational modes.[80] The bond strength changes according to the change in hybridisation state of the bond, with a decrease in bondstrength according to sp > sp^2 > sp^3. The force constant of the bond follows the same trend. The expected kinetic isotope effect is generally smaller for this type of isotope effect compared to bond breaking as the difference in force constant is smaller.

In transitioning from an sp^3 to an sp^2 centre it is the out of plane bending, which arises from reduced steric constraints around the sp^2 centre, that is mostly responsible for the observed isotope effect. The out-of-plane bending vibration observed in an infrared (IR) spectrum for a C–H bond on an sp^3 carbon is ~1350 cm^{-1}, while for a C–H bond on an sp^2 carbon is ~800 cm^{-1}. This leads to a theoretical secondary isotope effect for a C–H and a C–D bond of approximately 1.4. Typically, however, isotope effects of 1.1–1.2 are observed because the full conversion between sp^3 and sp^2 geometry is not present at the transition state. Clearly, such changes in hybridisation can also lead to inverse isotope effects, or isotope effects <1. When a reaction centre changes from sp^2 to a more sp^3-like state the difference in ZPE energy is greater at the transition state, leading to an inverse isotope effect.

β-secondary isotope effects arise from the electrons of a sigma bond which are involved in changes in hyperconjugation at the rate-determining step. An example is stabilisation of a cation through hyperconjugation, where the electrons of a sigma bond interact with the adjacent empty non-bonding p-orbital of the cation, this leads to weakening of the sigma

bond and results in a normal (>1) isotope effect. Inverse KIEs are possible through decrease in hyperconjugation (<1).

3.1.6 Techniques to measure kinetic isotope effects

There are three main methods used to determine the isotope effect of an enzyme-catalysed reaction. The first method is through the comparison of the reaction rates of isotopically labelled and unlabelled reactants while the other substrates are kept at a fixed concentration. This method is the only method that provides information on both catalytic turnover *(k_{cat})* and catalytic efficiency (k_{cat}/K_m), but is limited to larger isotope effects of at least 1.05.[59] A second, more sensitive, method to determine kinetic isotope effects is through isotopic perturbation of an equilibrium. This technique can be applied to systems with near degenerate equilibria and generates what is sometimes referred to as an equilibrium isotope effect (EIE). Kinetic isotope effects are measured as the rate of re-establishing chemical and isotopic equilibrium. This information is obtained by montoring the perturbation continuously.[83] The third method is an internal competition experiment. This is the most sensitive of the three methods. Here a reaction-sensitive position is measured against an internal control.[84,85] More specifically, the substrate of interest is isotopically substituted at, or directly adjacent to, the site of the reaction to report on the reaction mechanism. The labelled substrate is then either mixed with the naturally abundand isotopologue, when detection is by mass spectrometry, or with a second labelled substrate at a position remote from the reaction centre (typically at least two bonds away), when detection by radiation is used. The remote or unlabelled substrate act as an internal control for the unimpeded rate. The reaction is allowed to proceed to 20-70% conversion and the ratio of the isotopically labelled substrates is compared to the initial standard. The apparent isotope effect can then be found by comparing the ratio between the isotopically sensitive position and the internal control in the substrate (Equation 8) or in the product (Equation 9).[59,83] These equations hold for work with microscopic amounts of stable isotopic substrates, as well as for investigations where one of the competing species is present in radioactive tracer amounts only. In Equations 8 and 9, R_s and R_p are the isotopic ratio of substrate and product, respectively, at the observed fractional conversion, f. $^x(k_{cat}/K_m)$ gives the observed KIE, with x giving the mass of the heavier isotope. R_0 gives the isotopic ratio at the start of the reaction or at 100% completion. Examples of when this method is used is for

measuring ^{3}H- or ^{14}C-isotope effects, or when looking at natural abundance with an isotope-ratio mass spectrometer to determine changes in ^{13}C, ^{15}N, or ^{18}O mass ratio in CO_2 or N_2.

$$^{x}(k_{cat}/K_m) = \log(1-f)/\log(1-f)\,(R_s/R_0) \qquad \text{Equation 8}$$

$$^{x}(k_{cat}/K_m) = \log(1-f)/\log(1-f\,R_p/R_0) \qquad \text{Equation 9}$$

3.1.7 Commitment to catalysis

A major challenge when measuring KIEs for enzyme catalysed reactions can be forward and/or reverse commitment to catalysis (Figure 3.5).[2,86,87] When measuring isotope effects by internal competition, the measured or "apparent" KIE includes all steps from the enzyme and substrate(s) free in solution to the first irreversible step of the reaction coordinate. This includes non-chemical steps such as substrate binding and product release.[2] Commitment to catalysis can arise from any step whose rate constant appears in the expression for k_{cat}/K_m. Commitment to catalysis can lead to suppression or completely masking of the KIE on the reaction step. To overcome high commitment to catalysis a number of methods have been developed including measuring KIEs at pH values away from the optimum, through the use of alternative weakly binding substrates for forward commitment to catalysis (C_f) or alternative substrates that form unstable products for reverse commitment to catalysis (C_r).[60,61,85,88–91] Alternative weakly binding substrates overcome forward commitment by the increased dissociation constant.[92] This makes the concentration of enzyme bound species very low, leading to substrates only rarely undergoing transformation to the next enzyme bound species. The increased ligand dissociation rate reduces commitment to catalysis and the expression of the kinetic isotope effect on the reaction can be observed.[59,85,93] alternative substrates that form unstable products overcome reverse commitment by making the reaction the first irreversible step as the formed product rapidly decomposes after formation preventing the reverse reaction to occur.

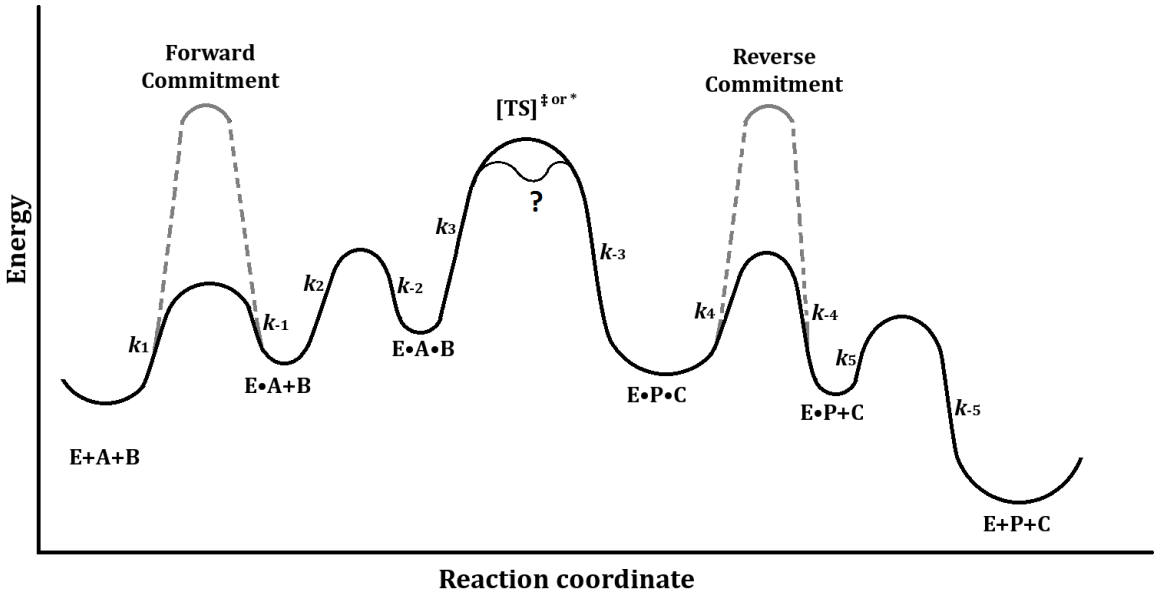


Figure 3.5. Hypothetical reaction coordinate diagram of a two substrate reactions.[85] *Dashed lines show a potential situation that could give rise to forward commitment, if after substrate binding the reaction is destined to go to product, or reverse commitment when product release is associated with a high energy step. When there is high commitment to catalysis (C_f or C_r) partly or completely suppressing the KIEs is observed.*

Intrinsic isotope effects can be obtained from the apparent isotope effect by correcting for the forward and reverse commitment using the expression derived by Northrop:[94]

$$\frac{k + C_f + C_r K_{eq}}{1 + C_f + C_r} = \left(\frac{k_{cat}}{K_m}\right) \qquad \text{Equation 10}$$

In this equation, k_{cat}/K_m is the observed isotope effect, C_f is the forward commitment to catalysis, C_r is the reverse commitment to catalysis, K_{eq} is the equilibrium isotope effect and k the intrinsic isotope effect. The expression can be reduced if the kinetic isotope effects are measured on a reaction that is irreversible. If one of the reaction products is effectively coupled off or the reaction is measured under initial velocity conditions this equation may sometimes also be used:

$$\frac{k + C_f}{1 + C_f} = \left(\frac{k_{cat}}{K_m}\right) \qquad \text{Equation 11}$$

3.2 Kinetic isotope effect measurements on ATP-PRTase

Goitein *et al.* studied a number of PRTase enzymes through kinetic isotope effect studies including ATP-PRTase from *Salmonella typhimurium*. However, in their work inconclusive results were obtained for the ATP-PRTase enzyme, with ^{14}C primary isotope effect of 1.01, an α-secondary isotope effect of 1.11 in the forward direction and isotope effects at unity in reverse. These results were attributed to commitment factors suppressing the observed isotope effect.[58] Based on these observations similar challenges were expected for the *M. tuberculosis*, *C. jejuni*, and *L. Lactis* ATP-PRTase enzymes. As previously discussed, commitment to catalysis can be overcome through the use of alternative substrates which are able to reduce or even fully suppress commitment.

In this chapter, the work towards the chemoenzymatic synthesis of isotopically labelled substrates, reaction conditions for measuring KIEs, measuring of commitment factors, measuring KIEs by internal competition, and the measurement of KIEs for *M. tuberculosis*, *C. jejuni* and *L. lactis* ATP-PRTase are discussed.

3.2.1 Selection of isotopic labels in KIE study

To study the reaction mechanisms by KIEs the most valuable information on the reaction is obtained from the primary and α-secondary isotope effects. It is these positions that undergo the biggest change going from reactants to the transition state. ATP-PRTase catalyses the condensation reaction between PRPP and ATP, whereby a new bond is formed between the anomeric carbon of PRPP and N1 of ATP, with inversion of stereochemistry at the anomeric carbon and accompanying loss of pyrophosphate (Figure 3.6).

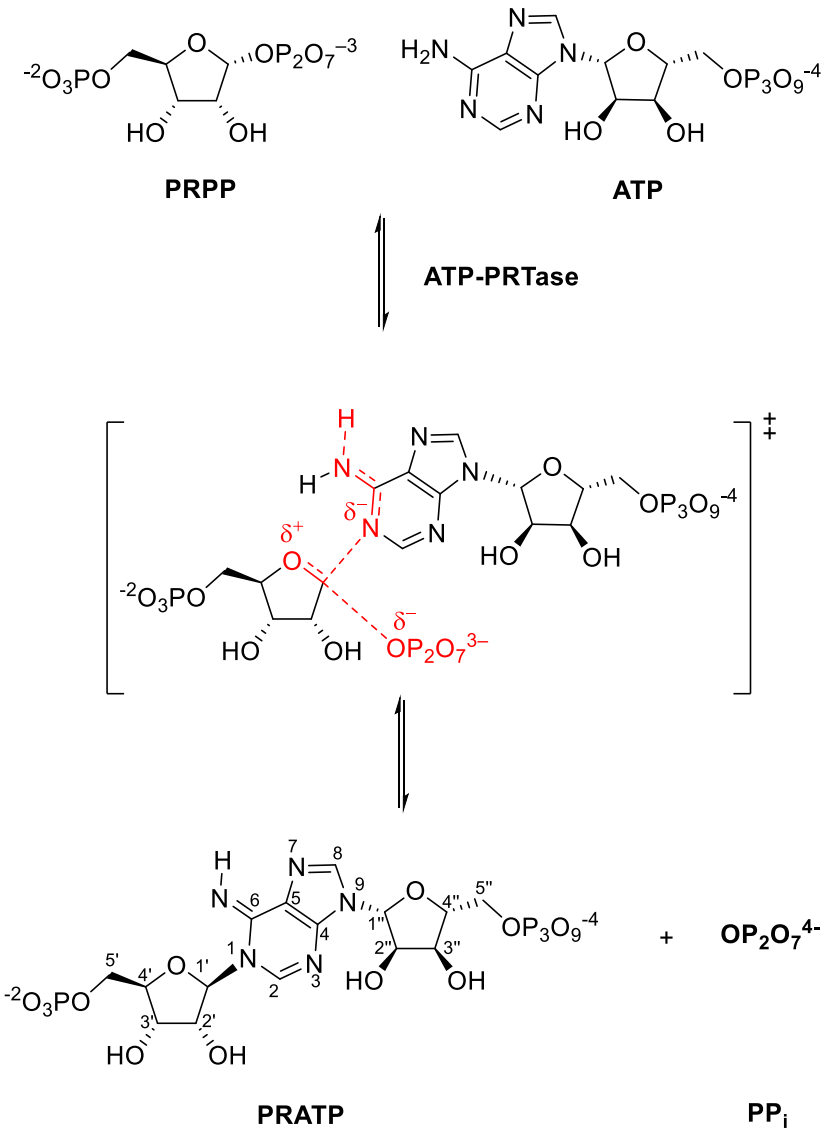


Figure 3.6. The reaction catalysed by ATP-PRTase. Highlighted in red are the atoms expected to undergo bond changes during catalysis.

In this work, studying the reverse reaction was preferred, as alternative substrates for pyrophosphate had been reported previously.[60,61] In addition the PRATP product was expected to be more stable than PRPP or PRPP analogues. The work from here on will be discussed in relation to the reverse direction unless stated otherwise.

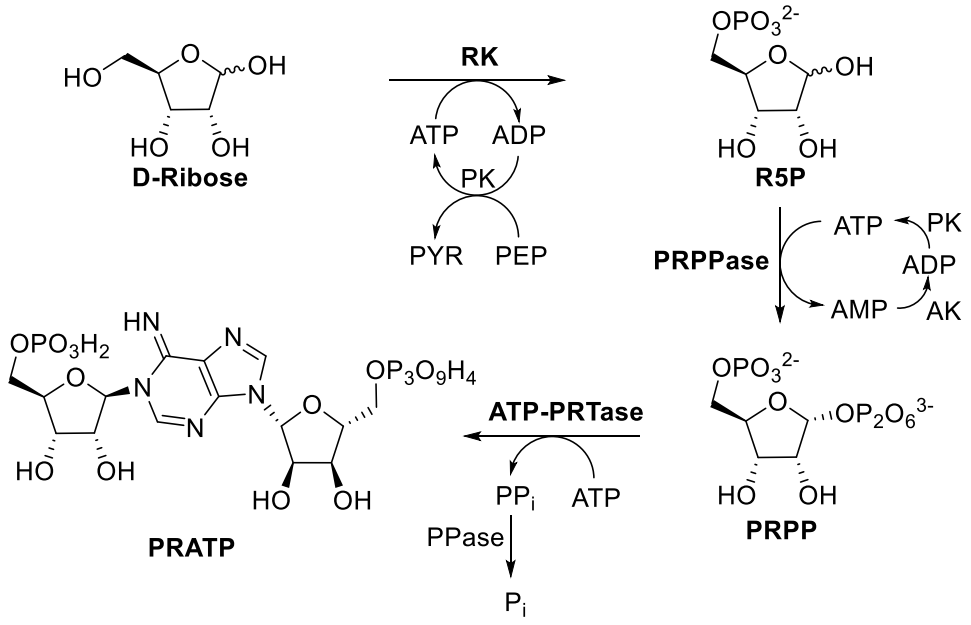
The primary isotope effect at the anomeric carbon (1'-C), will report on the bonding characteristics to the leaving group and incoming nucleophile of the activated species. As discussed in Section 3.1.5, the α-secondary position (1'-H) reports on hybridisation changes and will be important to support the findings for the primary carbon isotope effect. During

catalysis the bonding characteristics around the adenine ring of ATP are changed as well. The sigma bond is broken between the 1-N adenine nitrogen and the anomeric carbon affecting the conjugation of the aromatic ring. It is likely that the adenine 6-N will undergo a change in protonation state either before, during or after the reaction to allow for this change. These positions were thus expected to provide information on the change in conjugation and charge of the adenine ring. In addition, the 1-^{15}N KIE would give insight to the involvement of the adenine base at the transition state, information which is not necessarily clear from the primary and α-secondary isotope effects alone. The KIEs 2'-^{3}H or β-secondary and ribose ring oxygen, ^{18}O, provide additional information on the ribose ring conformation and the reaction mechanism; however, enough information was expected from the primary and α-secondary isotope effects and these sites were therefore not included in the study. Together this gave four sites of interest: 1'-C, 1'-H, 1-N, and 6-N. To establish the KIE of these sites the internal competition method was used (Section 3.1.6), with use of radioactive isotopes and detection by liquid scintillation counting. The isotopic substitutions are then ^{3}H, ^{14}C and ^{15}N for the relevant position. To measure the ^{15}N isotope effects a double isotopically labelled substrate was synthesised with the second label distal to the reaction site. In addition to the four sites of interest two internal standards were also included. To allow for the same method of separation during the KIE experiments the label was placed in the ribose 5-phosphate part of the molecule. The 5' position was chosen as it is at least two bonds away from the reaction site and thus expected to be unaffected by the reaction. This gave six isotopically labelled substrates that needed to be synthesised [1'-^{3}H]-, [1'-^{14}C]-, [1-^{15}N, 5'-^{14}C]-, [6-^{15}N, 5'-^{14}C]-, [5-^{3}H]- and [5-^{14}C]PRATP. The preparation of compounds containing these isotopic labels is discussed below.

3.2.2 Coupled enzymatic synthesis of phosphoribosyl adenosine triphosphate

Bearing in mind the synthesis of isotopically labelled substrates needed later in the project, and the amount of phosphoribosyl adenosine triphosphate (PRATP) needed in general, a coupled enzymatic synthetic route was considered for synthesising the unlabelled product. A method for the coupled enzymatic synthesis of PRATP has been reported.[58] This allows PRATP to be made in three enzymatic steps, starting from D-ribose (Scheme 3.1). Recycling of ATP prevents build-up of ADP and AMP in the reaction mixture, which are known inhibitors of the ATP-PRTase enzyme. It also limits the amount of ATP needed for the

reaction, which makes purification of the PRATP product after synthesis easier as well. The required adenylate kinase (AK) and pyruvate kinase (PK) were commercially available, whereas the ribose kinase (RK) and phosphoribosyl pyrophosphate synthase (PRPPase) needed to be expressed and purified. Plasmids bearing genes encoding for these two enzymes were kindly provided by Prof. V. L. Schramm (Albert Einstein College of Medicine, AECOM) along with the procedures for the expression and purification of the enzymes (Sections 8.3.1 and 8.3.2).[95,96]



Scheme 3.1 R5P: ribose 5-phosphate, Reactions were carried out in a phosphate buffer, in presence of NaCl and $MgCl_2$ pH 8.5. Substrates were added from stock solutions, pH adjusted to pH ~7.5.

Nuclear magnetic resonance (NMR) was used to obtain a more detailed picture of the reaction progress and the species present in the reaction mixture. This allowed for identification of problems, more challenging by other techniques such as UV analysis. The reactions were set up on a 1 mL scale and monitored by NMR employing PRESAT water suppression pulse sequence in presence of 15% D_2O (Figure 3.7).[97] After optimisation of buffer conditions, pH, enzyme concentration and metal concentration, good conversion (>70% by integration of the NMR peaks) of starting materials to product was achieved. After longer incubation times, the formation of the 6-*N*-PRATP side product (Dimroth rearrangement) was also observed. Formation of this side product can be prevented by

lowering the pH and keeping reaction times below 2 hours. Sufficient pure PRATP product can then be obtained by Biorad® P2 size exclusion gel filtration.

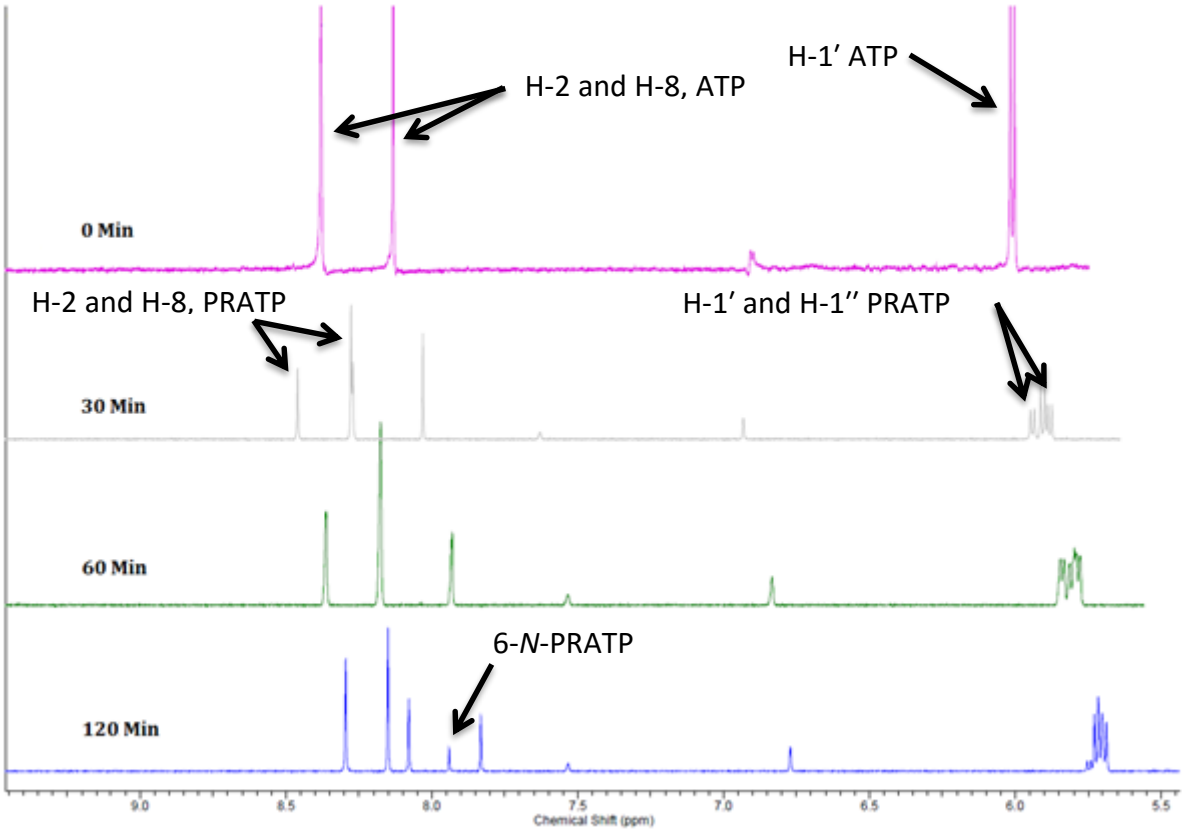
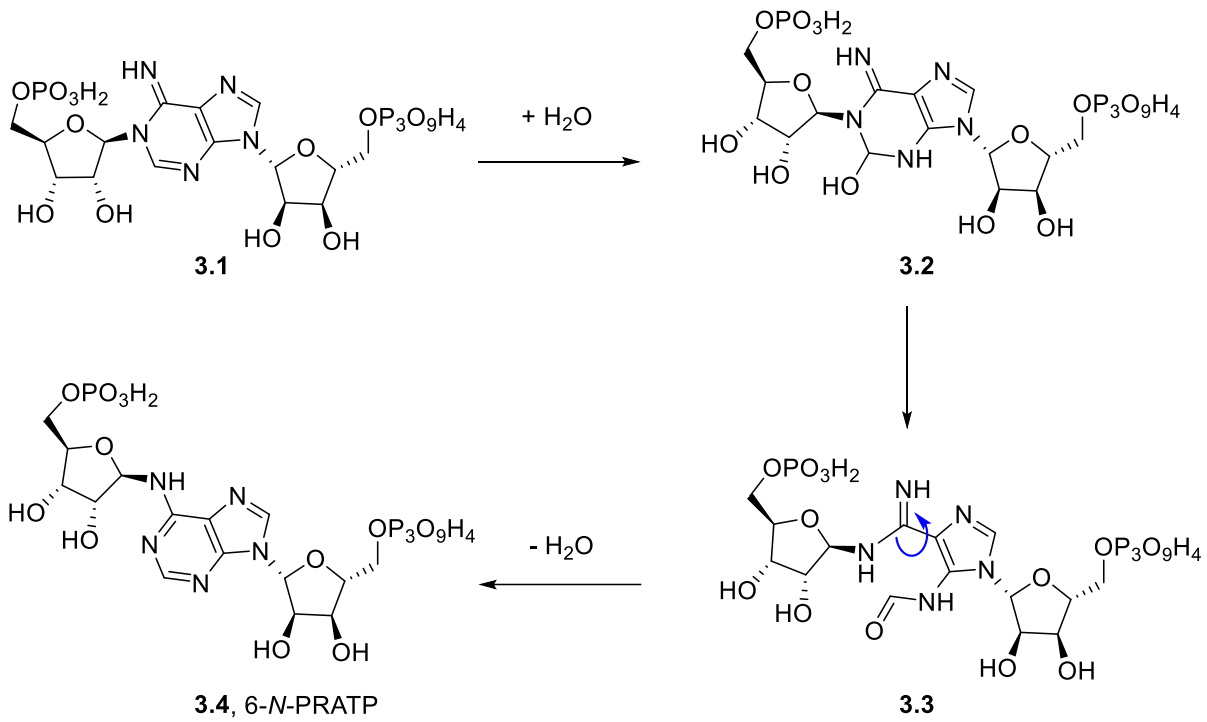


Figure 3.7. The coupled enzymatic synthesis of PRATP monitored by ^{1}H-NMR. With the reaction time in minutes, adenine protons of ATP are indicated. The change on the adenine ring is shown as two new peaks further downfield as well as the formation of two new anomeric protons. Characteristic peaks around the aromatic region of the rearranged side product are also shown.

3.2.2.1 Dimroth rearrangement

The first step of a Dimroth rearrangement is the addition of water to give the hemiiminal **3.2** followed by ring opening to give the aminoaldehyde product **3.3** (Scheme 3.2).[58,98,99] The ring opened product then allows rotation around the acetamidine and subsequent ring closure with release of water, restoring aromaticity to the adenine base leading to the 6-*N* substituted product **3.4**.



Scheme 3.2. The reaction mechanism of the Dimroth rearrangement leading to the 6-N substituted ATP side product. Blue arrow indicates rotation step leading to the rearranged side product.

3.2.3 Alternative substrates

The study by Goitein *et al.* showed high commitment to catalysis for *Salmonella typhimurium* ATP-PRTase, and so this was therefore also expected for the *M. tuberculosis*, *C. jejuni* and *L. lactis* ATP-PRTase enzymes.[58] An important aspect of this study was therefore to find conditions that would allow the measurement of KIEs. Previously, for the measurement of KIEs for orotate phosphoribosyltransferase (OPRTase), which catalyses similar chemistry to the ATP-PRTase enzyme, phosphonoacetic acid was successfully used to overcome commitment to catalysis.[60,61]

Phosphonoacetic acid was therefore tested for activity with the ATP-PRTase enzymes. Methyl bisphosphonate and fluoromethyl bisphosphonate were also tested based on their structural similarity to the substrate pyrophosphate (PP_i). The abilities of these PP_i analogues to act as substrates were tested spectroscopically and the relative rates of conversion compared to that of the natural substrate PP_i. Methyl bisphosphonate and fluoromethyl bisphosphonate gave similar rates to pyrophosphate under the conditions tested and thus expected to be unsuitable. A significant reduction in rate was observed for

the phosphonoacetic acid analogue (Tables 3.1). The catalytic efficiency was reduced 370 times for *C. jejuni* ATP-PRTase, and 500 times for *L. lactis*, compared to the natural substrate PP_i. The strong product inhibition for the *M. tuberculosis* enzyme made determining binding constants for this enzyme challenging. In addition, it was found that pyrophosphate started precipitating out before saturating conditions could be achieved (>2mM). Phosphonoacetic acid significantly reduced the rate of catalysis and, based on the findings of the other two enzymes, considered sufficient to continue.

It is important to note that a change in the transition state by the use of an alternative substrate cannot be ruled out, and examples of small alterations to the transition state by the use of alternative substrates have been reported.[61,62] However, high levels of detail on the transition state can still be obtained using this method and transition-state structures generated through the use of non-natural ligands have led to development of some of the most potent inhibitors known to date.[84] With transition-state guided inhibitor design being one of the goals of this research project, phosphonoacetic acid was considered to be a good candidate for KIE measurements and transition state analysis of ATP-PRTase.

Table 3.1. Kinetic parameters for ATP-PRTase *from C. jejuni and L. lactis with pyrophosphate (PP_i) and the alternative substrate phosphonoacetic acid (PA).*

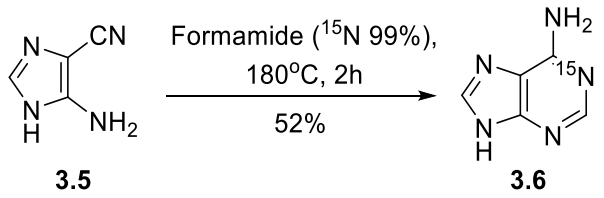
Enzyme	Substrate	K_m (μM)	k_{cat} (s^{-1})	Catalytic efficiency k_{cat}/K_m ($s^{-1}M^{-1}$)
C. jejuni	**PP_i**	96 ± 15	1.1 ± 0.05	11.5
ATP-PRTase	**PA**	157 ± 48	0.008 ± 0.001	0.051
L. lactis	**PP_i**	10 ± 3	0.84 ± 0.05	84
ATP-PRTase	**PA**	530 ± 145	0.089 ± 0.007	0.17

3.2.4 Isotopically labelled substrate synthesis

3.2.4.1 Synthesis of [1-^{15}N]adenine

The required ^{15}N-labelled substrates were not commercially available and therefore needed to be synthesised. Based on a modified literature procedure [1-^{15}N]adenine was expected to be accessible in a one-step synthesis from commercially available 5-aminoimidazole-4-carbonitrile (**3.5**) and ^{15}N labelled formamide (as in Scheme 3.3).[100,101] The challenge in this

synthesis was the small scale (<200µL total volume) and elevated reaction temperatures (180 °C) needed to drive the reaction. The higher temperatures led to partial evaporation of the formamide, concentrating the reaction and causing the starting material to precipitate out. After optimisation it was found that the best yield was obtained by performing the reaction in a closed off mass spectrometry vial and heating in an oil bath to ensure even heating under inert atmosphere. Following this procedure sufficient [1-^{15}N]adenine (**3.6**) was obtained for the isotope effect studies albeit in a modest 52% yield.

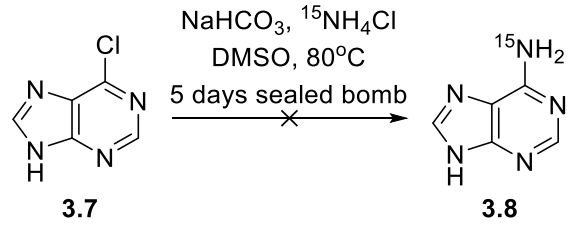


3.5 **3.6**

Scheme 3.3. Synthesis of [1-^{15}N]adenine with isotopically labelled formamide and 5-aminoimidazole-4-carbonitrile.

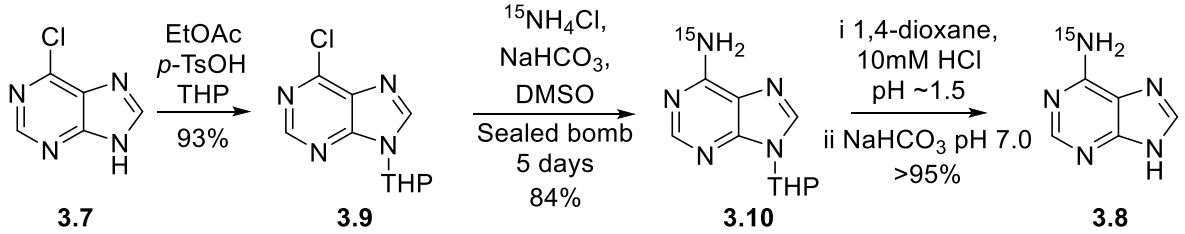
3.2.4.2 Synthesis of [6-^{15}N]adenine:

It was envisioned that [6-^{15}N]adenine would be accessible in one step through aromatic substitution of 6-chloropurine **3.7** and ammonia (Scheme 3.4). The preferred method was by McCann and Berti who used sodium carbonate and ammonium chloride in dimethyl sulfoxide (DMSO).[102] In this reaction, sodium carbonate reacts with ammonium chloride with the release of ammonia. It was found that the best yield could be obtained by performing the reaction in a sealed bomb for five days. The desired [6-^{15}N]adenine (**3.8**), was characterised by the splitting pattern of ~90 Hz of the amine proton. Side products were also formed. The major side product was identified as the dimerised product by Mass spectrometry. The similar nature of the side products to adenine made purification challenging and after several attempts by recrystallization, extraction and reverse phase chromatography this method was considered unpractical for the synthesis of [6-^{15}N]adenine and an alternative synthesis was investigated.



Scheme 3.4. One step synthesis of [6-^{15}N]adenine by nucleophilic aromatic substitution.

A three step synthesis of [6-^{15}N]adenine starting with 6-chloropurine and the incorporation of the tetrahydropyranyl ether (THP) protecting group has been successfully demonstrated by Milecky *et al.* (Scheme 3.5).[103] To this end, 6-chloropurine was subjected to 2,3-dihydro-2*H*-pyran and catalytic *p*-toluenesulphonic acid (*p*-TsOH) in ethyl acetate (EtOAc) to give protected purine (**3.9)** in 93% yield. In addition to masking the N-9 position the introduction of the THP protecting group also aids solubility in organic solvents which made purification of the intermediates easier. Nucleophilic aromatic substitution of protected purine (**3.9**) with ^{15}N labelled ammonium chloride and sodium carbonate in a sealed bomb then gave protected adenine (**3.10**) in an 84% yield. Acid hydrolysis followed by purification by DOWEX© H^+ exchange resin gave the desired [6-^{15}N]adenine (**3.8**)(95% yield).

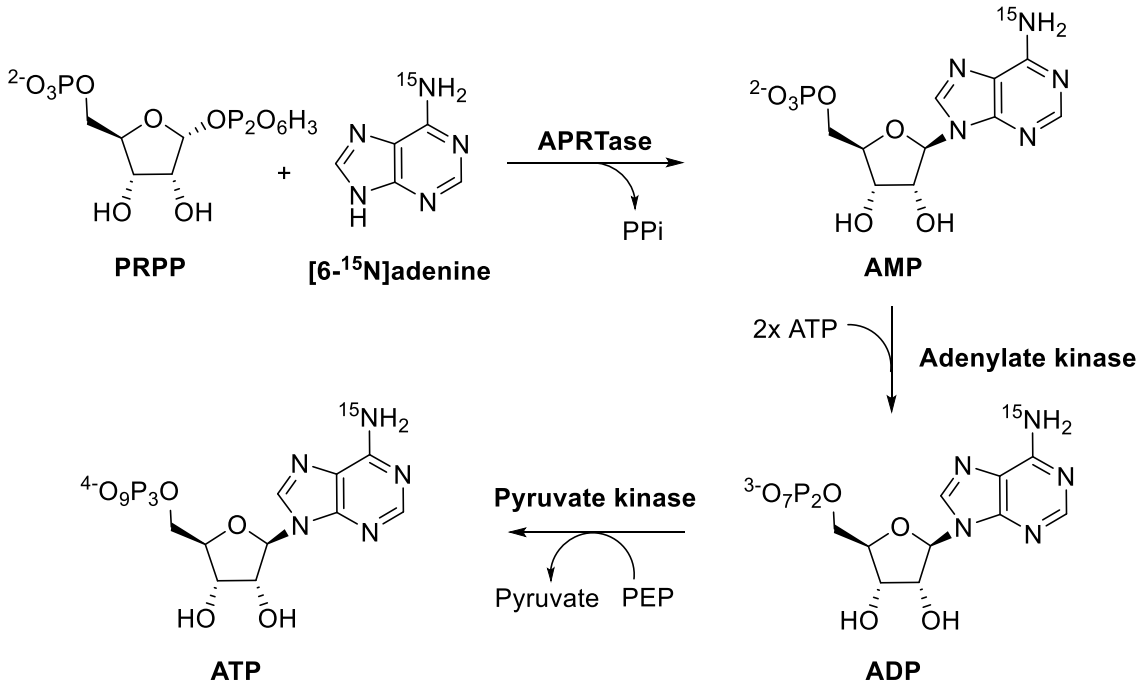


Scheme 3.5. Total synthesis of [6-^{15}N]adenine using the THP protecting group.

3.2.4.3 Synthesis of [1-^{15}N] and [6-^{15}N]ATP

Preparation of ^{15}N labelled ATP was readily achieved using a coupled enzymatic synthesis starting from PRPP and the appropriately ^{15}N labelled adenine (Scheme 3.6).[104] A small amount (<1%) of unlabelled ATP needed to be added to initiate the reaction. The reaction was then followed by HPLC (50 mM KH_2PO_4 Bu_4NHSO_4 pH 5.0, Section 8.3.7) and left at 37 °C until complete conversion of starting material to product was observed. Purification by semi-preparative HPLC using the same ion-pairing method, followed by desalting with 0.1%

trifluoroacetic acid in acetonitrile, gave the desired ^{15}N labelled ATPs. The products were concentrated by freeze drying, taken up in water and stored at -80 °C until further use.

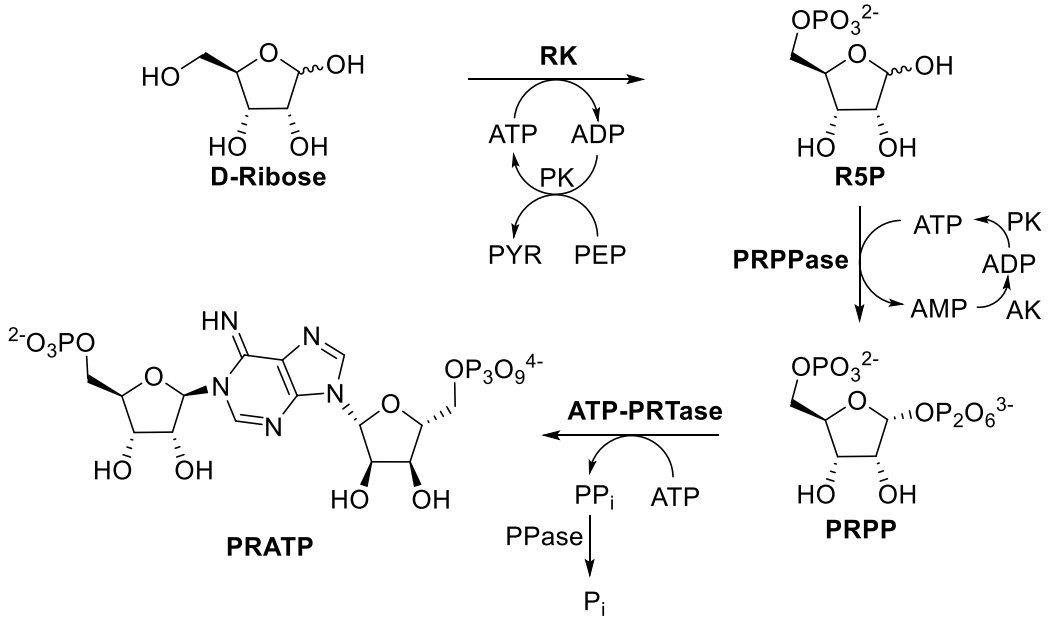


Scheme 3.6. Enzymatic synthesis of isotopically labelled ATP, example given for the synthesis of [6-^{15}N]ATP. APRTase: adenosine phosphoribosyl transferase, PRPP: phosphoribosyl pyrophosphate, PP_i: pyrophosphate, AMP: adenosine monophosphate, ADP: adenosine diphosphate, ATP: adenosine triphosphate, PEP: phosphoenol pyruvate. Reactions were carried out in a 50 mM phosphate buffer, in presence of NaCl 100 mM, $MgCl_2$ 10 mM, pH 8.5. Substrates were added from stock solutions pH adjusted to pH ~7.5.

3.2.4.4 Synthesis of isotopically labelled PRATP starting from ribose

The synthesis of [1'-^{3}H]PRATP and [1'-^{14}C]PRATP was achieved starting from the appropriately isotopically labelled D-ribose precursor. The synthesis of isotopically labelled PRATP starting from D-ribose was achieved following the conditions established for the synthesis of unlabelled PRATP (Scheme 3.7). Reactions were carried out using 20–50 µCi of isotopically labelled D-ribose. The reaction for the ^{3}H-labelled substrate was incubated for 30 min before additional unlabelled D-ribose was added. Unlabelled D-ribose was added to aid in manipulations following the reaction of this label. Reactions were followed by HPLC (reverse phase (C18), 50 mM KH_2PO_4 8mM Bu_4NHSO_4 pH 5.0, Section 8.3.7) and acidified after reaction completion or after two hours to prevent excessive formation of side products (Section 3.2.2.1). Isotopically labelled PRATP was then purified by HPLC (reverse phase (C18), 50 mM KH_2PO_4 8mM Bu_4NHSO_4 pH 5.0, Section 8.3.7), followed by drying using

speed vacuum evaporation. Samples were desalted shortly (at most one day) before KIE measurements were made using HPLC (reverse phase (C18), 50mM $AcONH_4$ pH 6.0 and acetonitrile, Section 8.3.7). Extreme care was taken to prevent even small amounts of degradation at this point, as this was found to greatly affect the KIE measurements. It was found that drying the samples for no longer than 10–15 min in small volumes (<200 µL) by speed vacuum evaporation gave the best results. After concentration the product was taken up in a small amount of water, divided into small aliquots, the radioactivity level was tested, and then the samples were frozen at -80 °C until further use. Each aliquot was used once and any remaining label was repurified before further use.

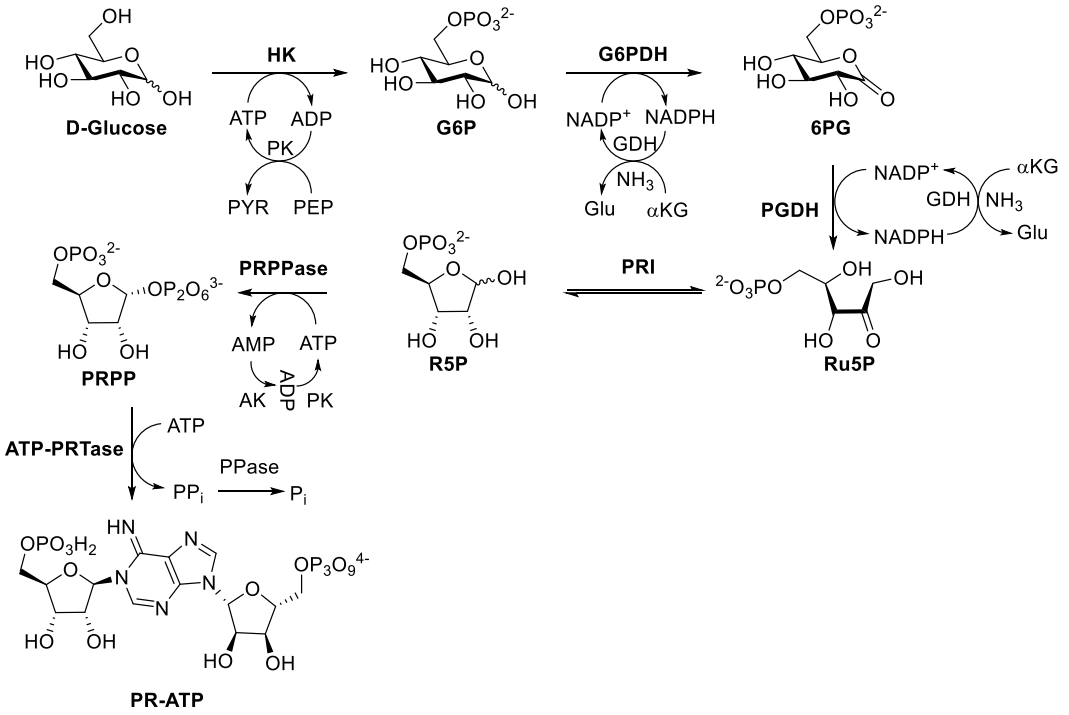


Scheme 3.7. RK: ribose kinase, PK: pyruvate kinase, R5P: ribose 5-phosphate, PRPPase: phosphoribosyl pyrophosphate synthase, AK: adenylate kinase; PRPP phosphoribosyl pyrophosphate, ATP-PRTase: ATP-phosphoribosyl transferase. Reactions were carried out in a phosphate buffer, in presence of NaCl, $MgCl_2$, Tris pH 8.5. Substrates were added form stock solutions, pH adjusted to pH ~7.5. Radiolabels were added, achieving radioactivity levels of 20–50 µCi.

3.2.4.5 Synthesis of isotopically labelled PRATP starting from glucose:

The synthesis of isotopically labelled PRATP starting from D-glucose comprises of six linear enzymatic conversions and three recycling steps and is a modified procedure for the synthesis of ATP published by Parkin *et al.* (Scheme 3.8).[96] The adenine PRTase (APRTase) enzyme was substituted for ATP-PRTase from *C. jejuni*. In addition, pyrophosphatase (PPase) was added, favouring the production of the PRATP product. Reactions were carried out on a 20–50 µCi scale. Unlabelled D-glucose was added to [6-^{3}H]-D-glucose after a 30 minute

incubation period because of the high specific activity (60 Ci/mmol) of this substrate. The double isotopically labelled PRATP reactions were carried out by exchanging ATP for the appropriately isotopically labelled counterpart. Again reactions were monitored by HPLC (Section 8.3.7) and left until no further conversion was observed or for two hours, whichever came first. The reaction was then purified by HPLC (Section 8.3.7). Overall the six isotopically labelled substrates required for this study were obtained starting either from D-ribose (Section 3.2.4.4) or D-glucose.



Scheme 3.8. Coupled enzymatic synthesis of isotopically labelled PRATP starting from d-glucose. HK, Hexo kinase; G6PDH, glucose 6-phosphate dehydrogenase; PGDH, Phosphogluconic acid dehydrogenase; PRI, Ribose 5-phosphate isomerase; PRPPase, phosphoribosyl pyrophosphate synthetase; PK, pyruvate kinase; GDH, l-glutamic dehydrogenase; AK, adenylate kinase; PEP, phosphoenolpyruvate; PYR, pyruvate; G6P, glucose 6-phosphate; 6PG, 6-phosphogluconate; Ru5P, ribulose 5-phosphate; R5P, ribose 5-phosphate; αKG, α-ketoglutarate; Glu, glutamic acid; NADP+/H, nicotinamide adenine dinucleotide phosphate. Reactions were carried out in a phosphate buffer, in presence of NaCl, $MgCl_2$, Tris pH 8.5. Substrates were added form stock solutions. Radiolabels were added, achieving radioactivity levels of 20-50 µCi.

3.3 Kinetic isotope effects for ATP-PRTases

3.3.1 Separation of substrate and product by charcoal/cellulose column

Isotopically labelled substrates and/or product are commonly separated either by HPLC or by solid phase extraction (SPE) using charcoal/cellulose columns.[86,104] The SPE columns have successfully been used for separating adenosine based substrates from ribose-like products. Separation of PRATP from the ribose reaction products ribose 5-phosphate (R5P) and/or

phosphoribosyl phosphonoacetic acid (PRPA) was therefore expected to be possible through this method. The advantage of using SPE columns is that it allows for faster processing of the samples and makes including multiple replicates easier. For these reasons separation by SPE was the preferred method and explored first.

To assess separation of the isotopically labelled starting material and product, conditions needed to be found which would retain the isotopically labelled PRATP but would also elute reaction products. For this, 0% and 100% conversion were chosen as the positive and negative controls. To favour the reaction equilibrium towards the PRPA and ATP products, a large access of phosphonoacetic acid was added. An excess of the nucleophile was expected to push the equilibrium towards the PRPA and ATP product. To allow for a direct comparison, a master mix containing isotopically labelled PRATP and phosphonoacetic acid in Tris buffer (50 mM Tris, 100 mM NaCl, 50 mM KCl, 10 mM $MgCl_2$, pH 8.5) was divided into two. To achieve 100% conversion an excess of *C. jejuni* enzyme was added to one part and buffer to the other part. These were split another time, one to act as a reference count and one to test retention of the label onto a charcoal column.

After optimisation it was found that addition of ammonium carbonate (pH 6.0) facilitated the separation of the PRATP substrate from the ribose products. In addition, the reaction was confirmed to go to completion under the used reaction conditions described above with the small discrepancy ascribed to pipetting errors (Table 3.3).

Table 3.2. The results from the charcoal retention experiment and selective elution of isotopically labelled ribose recovery after 100% conversion. Results are from a 50 mM ammonium carbonate pH 6.0 elution buffer.

	No enzyme	**100% Completion**
Charcoal	240 cpm	21900 cpm
Reference	22000 cpm	21600 cpm

3.3.2 Commitment to catalysis by isotope trapping

The forward commitment to catalysis, or the partitioning of the Michaelis complex to product relative to substrate release, was measured by the isotope trapping method.[85] The experiment determines the fraction of enzyme-bound isotopically labelled PRATP that

converts to product after a large excess of unlabelled PRATP and the second substrate PA is added. Fractions are taken after one and several turnovers and extrapolated back to find the initial amount of substrate released. The commitment factor can be found through Equations 12 - 14.[85] E_t represents the concentration of enzyme active sites, ^{14}CPRATP gives the concentration of substrate, K_m is the michaelis constant for PRATP and [*EA*] gives the concentration of enzyme-bound isotopically labelled PRATP. [PRPA] represents the initial amount of product formed of a large excess of unlabelled PRATP and phosphonoacetic acid is added.*Y* gives the fraction of product over the concentration of enzyme-bound ^{14}C-PRATP. C_f is the forward commitment factor. An example calculation can be found in Section 8.3.9.

$$\frac{[E_t][^{14}\mathrm{CPRATP}]}{K_\mathrm{m} + [^{14}\mathrm{CPRATP}]} = [EA] \qquad \text{Equation 12}$$

$$\frac{[\mathrm{PRPA}]}{[EA]} = Y \qquad \text{Equation 13}$$

$$\frac{Y}{1-Y} = C_f \qquad \text{Equation 14}$$

The *C. jejuni* enzyme showed low commitment to catalysis with a commitment factor of 0.030 ± 0.004 (Table 3.4). A similar commitment factor was obtained for the *L. lactis* ATP-PRTase enzyme complex with a commitment to catalysis of 0.028 ± 0.004. Commitment factors of similar magnitude were also observed for *H. sapiens* and *P. falciparum* OPRTase at 0.042± 0.006 and 0.035 ± 0.003.[61] This suggests that commitment to catalysis should not affect the KIE measurements for these enzymes.

No commitment to catalysis could be observed for *M. tuberculosis* ATP-PRTase; however, in addition no reaction was observed under the conditions used. The isotope trapping experiment relies on the addition of a large excess of unlabelled PRATP (500 µM) to outcompete the isotopically labelled substrate for binding back to the enzyme after the initial incubation period. The *M. tuberculosis* ATP-PRTase enzyme shows strong inhibition by PRATP significantly reducing the observed rate. How product inhibition affects the experiment is unknown. However, based on the reduced rate by phosphonoacetic acid, similarly to the other two enzymes studied, and the observation of low commitment to

catalysis for the *C. jejuni* and *L. lactis* ATP-PRTase enzyme, suppression of KIEs for *M. tuberculosis* ATP-PRTase was not expected.

Table 3.3. Forward commitment factors for C. jejuni and L. lactis ATP-PRTase.

Enzyme	Forward commitment (C_f)
***C. jejuni* ATP-PRTase**	0.030 ± 0.004
***L. lactis* ATP-PRTase**	0.028 ± 0.004
***M. tuberculosis* ATP-PRTase**	-

Error given as standard deviation in measured PRPA concentration.

3.3.3 Measuring kinetic isotope effects for ATP-PRTase

In all the KIE experiments performed, a master mix containing all components except for the enzyme was made (Section 8.3.8). The master mix was then divided into seven equal volumes to give four reactions for partial conversions (20% to 40%), two reactions for 100% conversions and a negative control. The reactions were incubated at 25 °C for 10 to 25 min depending on the enzyme. The reaction was acidified with sulphuric acid to pH 6.0 to stop the reaction and then directly loaded onto pre-equilibrated charcoal columns and the column washed with elution buffer. The eluent was collected in scintillation vials, and dried by speed vacuum evaporation for at least 12 hours. The resulting solids were taken up in 500 µL of water and 10 mL scintillation fluid was added. The samples were then counted on a scintillation counter for 10 cycles of 10 min per sample. Channel A was set for ^{3}H (0.0–25.0 keV and channel B for the remaining ^{14}C signal (25.0–400.0 keV). In each sample set, a ^{14}C standard was included to correct for channel overlap between the two isotopes. The isotope effects were then calculated from the average cpm/isotope by Equation 9 (Section 3.1.6). For a more detailed description see Section 8.3.8, which includes an example calculation. Each position was measured for at least two independent experiments with a minimum of eight replicates in total. Reliability of the results is reported as standard error of the mean (SEM).

3.3.3.1 KIE results for *Mycobacterium tuberculosis* ATP-PRTase

The primary isotope effect 1'-^{14}C reports directly on the nature of the nucleophilic substitution at the transition state. This isotope effect provides information on the

interaction of the anomeric carbon and is sensitive to the motion along the reaction coordinate of this position during catalysis. KIEs between 1.01–1.08 suggest weak interactions with partial involvement of the nucleophile and/or leaving group at the transition state.[60,104] If an isotope effect of around unity is found, that indicates a fully dissociated transition state with a carbocation intermediate. Values above 1.08 indicate an A_ND_N or S_N2 type reaction mechanism.[62,105,106] Large isotope effects (>1.25) for the α-secondary KIE 1′-^{3}H are characteristic of a dissociative reaction mechanism with limited participation of the nucleophile and leaving group at the transition state.[102]

The *M. tuberculosis* ATP-PRTase enzyme gave a primary isotope effect of 1.028 ± 0.001 1′-^{14}C which suggests weak bonding interactions at the transition state to the nucleophile and/or leaving group (Table 3.5). This is supported by the large 1′-^{3}H isotope effect of 1.250 ± 0.008 which suggest a well-developed ribocation with double bond characteristics between the ribose ring oxygen and the anomeric carbon. An α-secondary isotope effect of 1.250 ± 0.008 is close to the maximum isotope effect reported for dAMP hydrolysis in water further supporting that commitment to catalysis is not affecting the isotope effect values and the measured isotope effects are intrinsic.[102] The dissociative mechanism is further supported by the large 1-^{15}N isotope effect of 1.024 ± 0.001 which approaches the theoretical maximum of 1.04 for this isotope label.[87] The 1.012 ± 0.001 6-^{15}N isotope effect indicates considerable change in conformational freedom around the 6-*N* amine of ATP.

Table 3.4. KIE results for Mycobacterium tuberculosis ATP-PRTase.

Position	KIE	SEM	n
1′-^{3}H	1.250	0.008	8
1′-^{14}C	1.028	0.001	8
1-^{15}N, 5′-^{14}C	1.024	0.001	12
6-^{15}N, 5′-^{14}C	1.012	0.001	8
5-^{3}H	1.017	0.001	8

3.3.3.2 KIE results for *Campylobacter jejuni* ATP-PRTase

The intrinsic KIEs for ATP-PRTase from *C. jejuni* suggest that the reaction goes through a dissociative D_N*A_N reaction mechanism based on the primary isotope effect of 1.029 ± 0.002 (Table 3.6). The large α-secondary isotope effect of 1.212 ± 0.002 further supports these findings. The isotope effect of near unity for the primary 1-*N* was surprising given the large isotope effects found for the other two ATP-PRTase enzymes.

Table 3.5. KIE results for Campylobacter jejuni ATP-PRTase.

Position	KIE	SEM	n
1′-^{3}H	1.212	0.002	10
1′-^{14}C	1.029	0.002	8
1-^{15}N, 5′-^{14}C	0.995	0.003	8
6-^{15}N, 5′-^{14}C	1.005	0.001	8
5-^{3}H	1.009	0.001	8

3.3.3.3 KIE results for *Lactococcus lactis* ATP-PRTase

The 1′-14C intrinsic KIE for *L. lactis* suggests that this short form of the ATP-PRTase enzyme also goes through a D_N*A_N reaction mechanism with an unstable carbocation intermediate (Table 3.7). The α-secondary of 1.147 ± 0.006 supports these findings, although given the small primary 1′-^{14}C and large 1-^{15}N isotope effect, a larger 1′-3H isotope effect was expected. ^{3}H isotope effects are also sensitive to other interactions at the transition state, such as H-bonding and van der Waals interactions, potentially reducing the conformational freedom and partly suppressing the α-secondary isotope effect. The 1-^{15}N isotope effect at

the theoretical maximum suggests a fully dissociated adenine base with only interactions between the phopshonoacetic acid and the ribose ring remaining at the transition state. The isotope effect at unity for 6-^{15}N suggests changes at this position are not involved at the transition state.

Table 3.6. KIE results for Lactococcus lactis ATP-PRTase.

Position	KIE	SEM	n
1′-^{3}H	1.147	0.006	10
1′-^{14}C	1.031	0.001	8
1-^{15}N, 5′-^{14}C	1.028	0.001	8
6-^{15}N, 5′-^{14}C	0.996	0.003	8
5-^{3}H	1.003	0.001	8

Chapter 4

Transition state modelling

4.1 Introduction

The modelling of transition states through density functional theory (DFT) calculations, coupled with experimental kinetic isotope effects, can provide a detailed representation of the reactive specie involved at the rate determining step. To this end, the transition states of the reaction catalysed by ATP-PRTase were investigated by DFT with the experimental KIEs forming the boundary conditions for the transition state model. All optimisations were carried out *in vacuo*. A structure was considered optimised when a calculation successfully converged and only one imaginary frequency corresponding to the reaction coordinate was found. The calculated isotope effects were generated by comparing the ground state (GS) to the transition state with the ISOEFF98 program.[107] ISOEFF98 is a program that uses the force constants from the GS and the TS to calculate the corresponding isotope effect.[107,108] The program includes a scaling factor optimisation, which allows correction of the vibrational frequencies for the used functional and basis set.[107,109–111] The refinement process of the GS and TS will be discussed below.

4.2 IUPAC nomenclature to describe enzymatic nucleophilic displacement and substitution reactions

In this work the International Union of Pure and Applied Chemistry (IUPAC) nomenclature will be used to describe enzymatic nucleophilic displacement and substitution reactions.[105] Enzyme catalysed reactions differ from reactions in solution, in that enzyme active sites can introduce additional geometric restrictions upon a reaction. In an A_ND_N, or an "S_N2", reaction the first step involves association of a nucleophile to the reaction centre. This leads to a degree of symmetry at the transition state, with bond lengths to both the association nucleophile (A_N) and the dissociation nucleophile (D_N) smaller than 2.0 Å and overall inversion of the stereochemistry in the product (Figure 4.1). For D_N*A_N or "S_N1-like" reactions dissociation of the leaving group occurs first, followed by association of the nucleophile. The reaction may go through an unstable intermediate (*) or a transition state (‡) depending on its lifetime, the short lifetime precludes chemical detection or isolation. Reactions with a stable intermediate (+) is a D_N+A_N reaction.[112] When the nucleophile commences its approach (~ 3.0 Å) prior to full dissociation of the leaving group (<2.0 Å), this is referred to as a D_NA_N reaction.

As per Chapter 3, when discussing the ATP-PRTase catalysed reaction throughout this chapter the reaction steps are discussed in reverse order. When the physiological direction is discussed this will be specified.

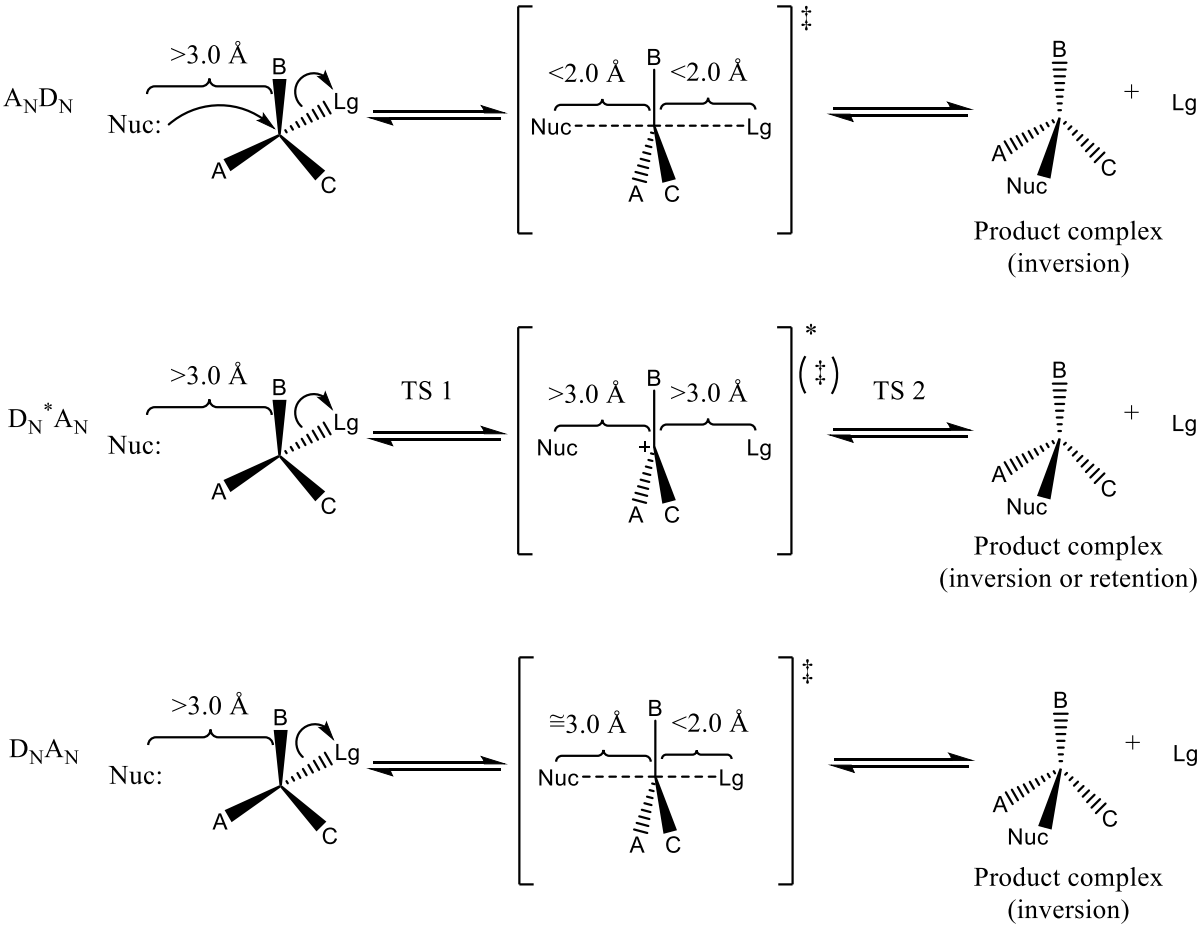


*Figure 4.1. Different types of nucleophilic displacement and substitution reactions for enzyme catalysed reaction following the IUPAC nomenclature. Nuc, Nucleophile; Lg, leaving group; D_N, dissociation nucleophile; AN, association nucleophile; *, unstable intermediate; ‡, transition state; TS, transition state. Distances in transition states represent the physical separation that needs to occur for this type of transition state to form. Image is a modified figure taken from Schramm and Shi.*[112]

4.3 The ground state model

Experimental KIEs represent the isotope effect from free in solution to the highest energy barrier of the reaction pathway, the transition state. A ground state model used in calculating the kinetic isotope effect should therefore best represent the lowest energy state of the substrate in solution. As the input model, crystal structures can provide a good starting point for the geometry optimisations. To ensure that the full conformational space

will be sampled, different starting points are generated. Refinement of these structures was employed using Gaussian 09, Gaussian, Inc.[113]

PRATP bound to *C. jejuni* ATP-PRTase core mutant was used as the input structure.[67] Large structures such as PRATP can make optimisation computationally time consuming; the model was therefore simplified to include the ribose-5-phosphate adenine in the ground state structure (Figure 4.2). Generally atoms greater than two bonds away from the reaction centre are unlikely to contribute significantly to the kinetic isotope effect.[114,115] Therefore, truncation of the model was not expected to affect the calculated isotope values around the reaction centre. The 5'-phosphate was considered deprotonated based on the experimental conditions (pH 8.5) and reported pK_a for related structures.[116]

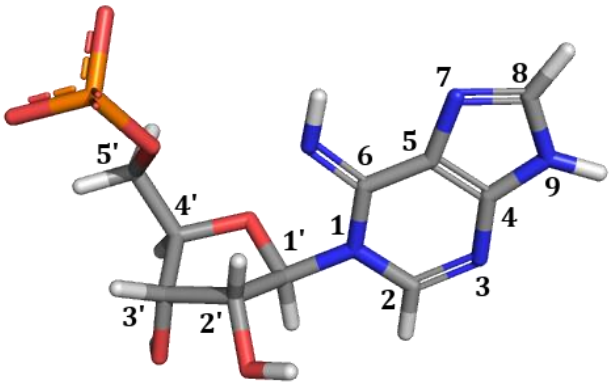


Figure 4.2. The truncated ground state structure *generated from the PRATP bound C. jejuni ATP-PRTase core mutant crystal structure.*[67]

Seven different starting conformations were generated by changing the dihedral angle of bonds that could rotate freely. These input structures were optimised using the B3LYP and M06-2X functionals and the 6-31G(d) basis set.[117–119] The Becke three parameter exchange functional in combination with the Lee-Yang-Parr correlation functional (B3LYP) has successfully been used for the generation of a number of transition states for reactions involving D-ribose.[61,62,84,85] B3LYP offers a high level of accuracy across different systems at reasonable computational cost. The M06-2X functional is designed for a number of applications, including excited states such as transition states and was therefore also employed.[107,118] A structure was considered optimised when the calculation successfully met the convergence criteria and there were no imaginary frequencies for the optimised

structure. The energy (kcal/mol) of the structure was then compared for each functional to find the lowest energy state (Table 4.1).

Table 4.1. Energy of the three lowest energy ground state structures after geometry optimisation by DFT.

Functional	Structure	Overall energy (kcal/mol)	Δ (kcal/mol)
B3LYP	**1**	-960085.0	0
	2	-960077.9	7
	3	-960072.2	13
M06-2X	**1**	-959800.8	0
	2	-959794.0	7
	3	-959790.5	10

An interesting trend was observed for the optimised ground states. The position of the imine proton (6-NH) greatly affected the conformation of the adenine base. When the proton points towards the ribose ring, an interaction was formed with the 5'-O and the 2-H or 6-NH proton (Figure 4.3, **A** and **B**) whereas for structures with the 6-NH proton pointing away from the ribose ring a structure with a hydrogen bond (H-bonds) between 6-N and 2'-H was usually found (Figure 4.3, **C**).

The lowest energy structures were the same for both the B3LYP and M06-2X functionals (Figure 4.3, **A** and **B**). Structure **A** and **B** both form four H-bonds, two of which are conserved in both structures. These are the H-bond between 3'-H and O-phosphate and the H-bond between 3'-OH and 2'-O. In structure **A** there is a H-bonding interaction between the 2'-H and the oxygen of the phosphate group at an angle of 175°. Another interaction is made with the same phosphate oxygen and the 2-H of the adenine ring at an angle of 172°, which is close to the optimum of 180° for H-bonds.[120] The ribose ring adopts the slightly less favourable 3-exo ring pucker, potentially resulting from interactions between the phosphate oxygen and the 2-H proton. In structure **B,** the adenine ring position allows for a H-bond to form between the imine proton and the 5'-oxygen at an angle of 168°. In addition, an interaction is made between 2'-OH and the phosphate group at an angle of 145°. The ribose ring pucker adopts the favourable 2-endo conformation in this structure. Overall, for

structure **A** having the H-bond closer to the optimum bond angle of 180° outweighs the slightly less favourable ribose ring pucker and leads to the lowest energy ground state structure. Structure **A** was thus used for comparison to the transition state models in the rest of this study (Section 8.4.3).

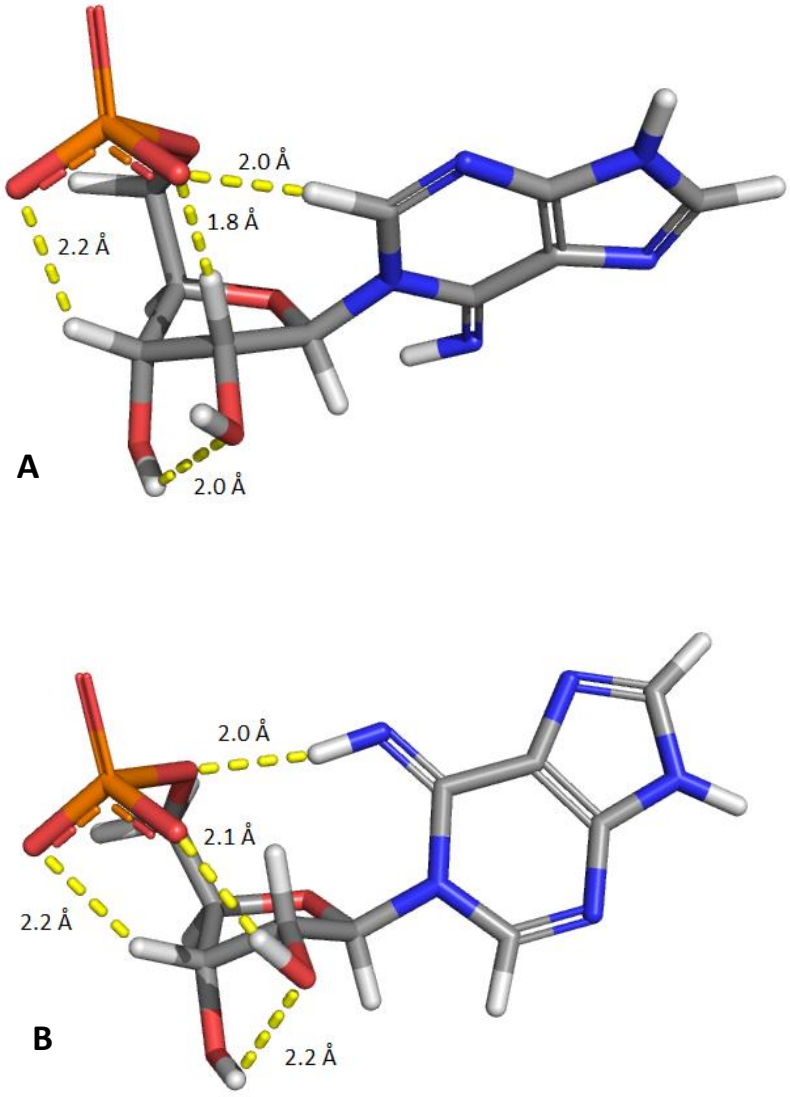


*Figure 4.3. GS models **A** the lowest energy structure for both functional, this GS model was used to generate calculated KIEs. **B** the ground state structure that ranked second lowest overall, which was also the same for both functionals .*

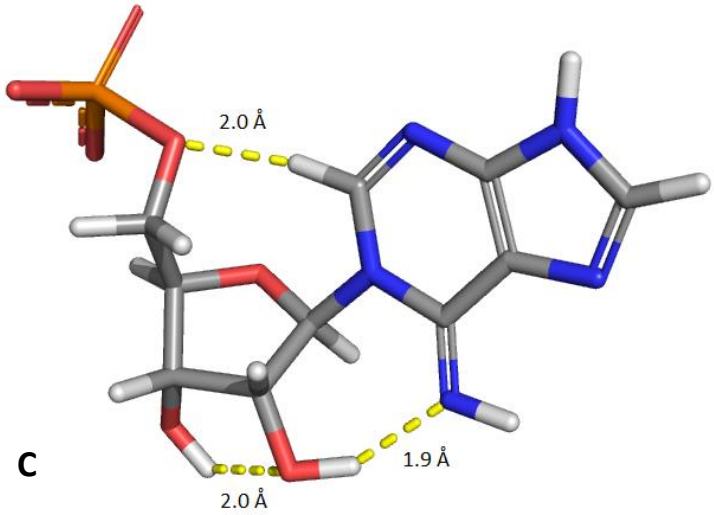


Figure 4.3. (continued). ***C*** *is an example of the effect of the change in proton position on the imine group resulting almost exclusively in structures with the adenine base in this position.*

4.4 Molecular electrostatic potential surface

The molecular electrostatic potential surface (MEPS) is a means to visualise the characteristic molecular electrostatic potential, displayed on a van der Waals surface. [121,122] The MEPS provides a template for inhibitor design by highlighting where charges are built up.

The MEPS of the ground state structure is represented in Figure 4.4 and shows a generally neutral ribose part with the more electron rich ribose ring oxygen. On the adenine there is charge build up around the 1-N, 3-N and 7-N nitrogen's and overall a neutral aromatic ring system. The electron rich nature of the phosphate group is highlighted by the negative charge of this position.

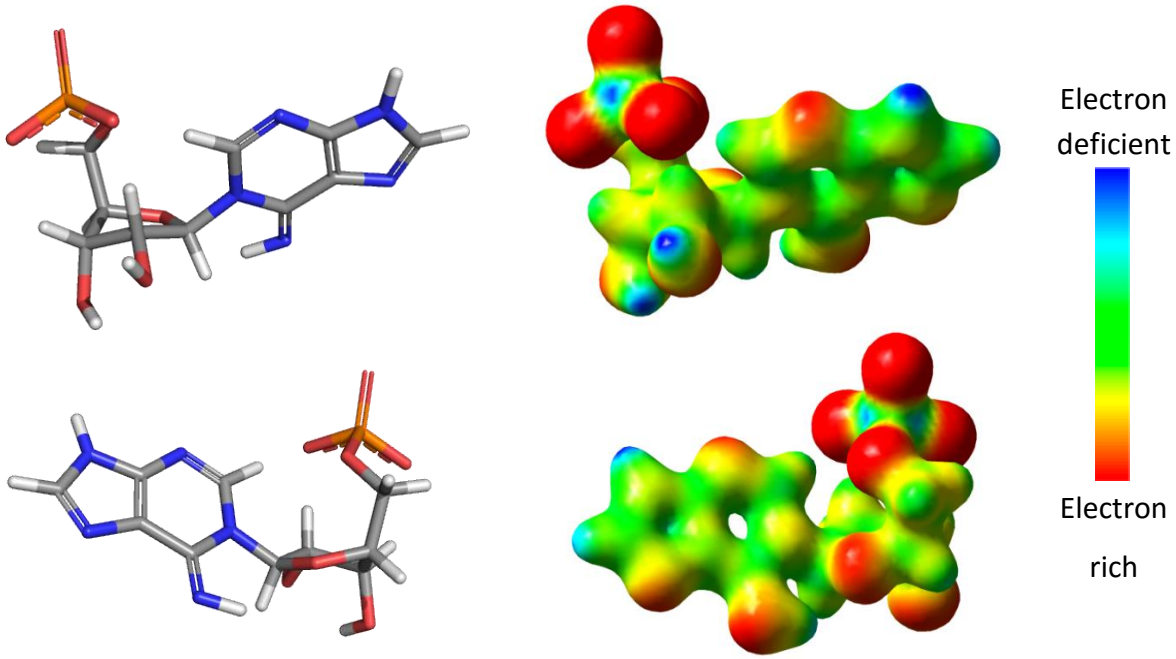


Figure 4.4. Molecular electrostatic potential surface (MEPS) of the ground state structure used in this study. The MEPS are shown on the right hand side at two different orientations, with the images on the left showing the corresponding stick representation. In the MEPS, red indicates electron rich or a relative negative charge, green refers to a neutral position and relatively positive atoms are shown in blue.

4.5 The transition state model

The experimental primary 1′-^{14}C KIE of 1.028-1.031 and α-secondary 1′-^{3}H isotope effects between 1.147–1.250 suggest a dissociative reaction mechanism for the ATP-PRTase enzymes with low bonding order to the leaving group and/or the incoming nucleophile.[105] The 1-^{15}N nitrogen isotope effect further supports this mechanism. Work towards the generation of the transition state model and matching to the experimental values introduced in Section 3.3.3, will be discussed below.

4.5.1 Adenine model to match the nitrogen isotope effect

The 1-^{15}N and 6-^{15}N KIEs arise from the extent of C1′-N1 bond dissociation as well as leaving group interaction at the transition state (Section 3.2.1). A positive KIE indicates more conformational freedom at the transition state and an inverse KIE suggests a more constrained environment. The theoretical magnitude of the ^{15}N isotope effect is 1.04 with values ranging from 1.010 to 1.037 reported for reactions involving aromatic nitrogenous bases.[123,124] In most purine *N*-ribosyl transferases, protonation of the purine at N7 is proposed as part of the transition state structure.[104,125] Protonation at this position makes

the adenine electron deficient, weakening the *N*-ribosidic bond and facilitating the reaction by forming a neutral leaving group. Overall the transition state was generally matched to the primary ^{14}C isotope effect and small differences in the primary ^{15}N were accepted as being within experimental uncertainties. In some cases the difference was ascribed to not considering the environment around the leaving group in the model, such as π-stacking, H-bonding and hydrophobic interactions.[85,104]

M. tuberculosis and *L. lactis* ATP-PRTase gave large primary nitrogen isotope effects within the range of a fully dissociated purine ring, when compared to previously reported transition state structures. A nitrogen isotope effect at unity was found for *C. jejuni* ATP-PRTase. The kinetic isotope effect of a fully dissociated adenine was thus investigated first. The effects of charge, tautomer state and the position of imine proton in either the *E* or *Z* configuration were also examined. This work was carried out at the B3LYP/6-31^{+}G(d,p) and M06-2X/6-31^{+}G(d,p) level. Different functionals were chosen to investigate their performance in modelling the transition state. The B3LYP functional is most widely used in transition state model calculations and therefore included.[61,85,126] The M06-2X is upcoming and has been used successful for other systems.[127,128] The results of the two functionals can be found in Table 4.2 and Figure 4.5.

Table 4.2. The effect of protonation and tautomeric state of the adenine ring on the calculated 1-^{15}N isotope effect.

	B3LYP		**M06-2X**	
State	**1-^{15}N**	**6-^{15}N**	**1-^{15}N**	**6-^{15}N**
Neutral	1.016	0.989	1.022	0.994
Anionic *E*	1.023	1.004	1.026	1.007
Anionic *Z*	1.025	1.004	1.027	1.007
Neutral N3 tautomer *E*	1.019	1.000	1.020	1.005
Neutral N3 tautomer *Z*	1.022	0.999	1.023	1.006
Neutral N7 tautomer *E*	1.022	1.007	1.022	1.010
Neutral N7 tautomer *Z*	1.023	1.009	1.023	1.013
Protonated at N3	1.014	0.988	1.018	0.988
Protonated at N7	1.017	0.996	1.021	0.991

- not calculated for this functional

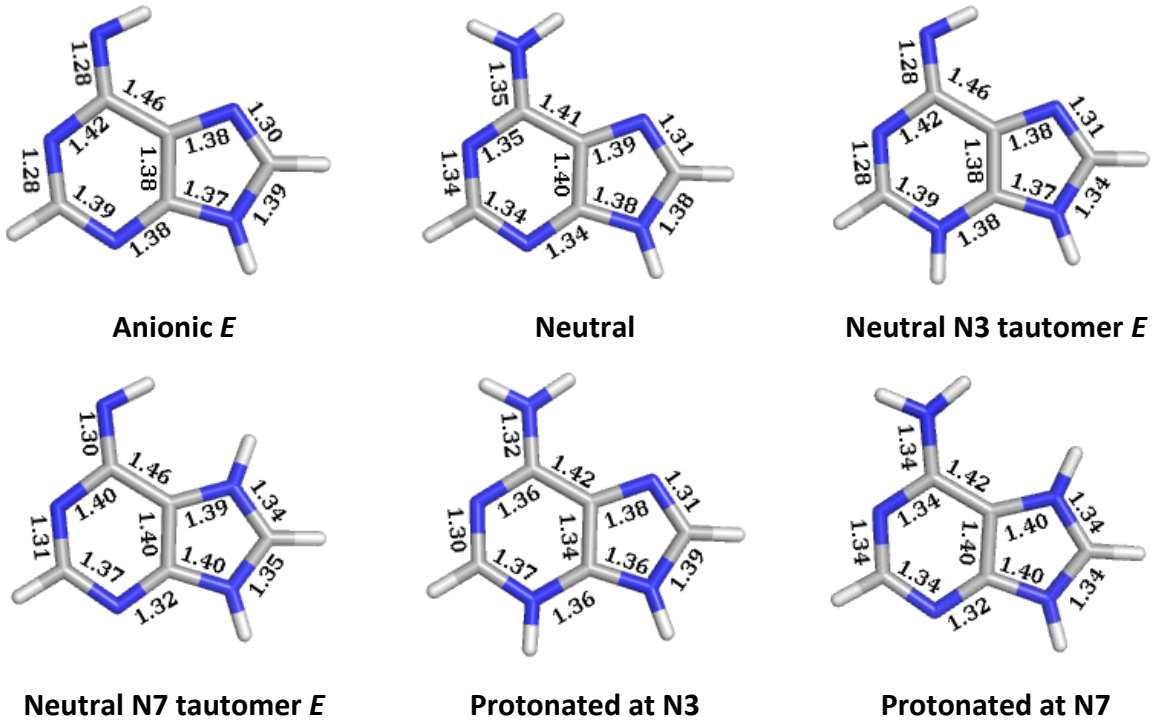


Figure 4.5. Different adenine charge and tautomeric states. Bond lengths are given in Å next to individual bonds.

The charge, imine proton position and tautomer state of the adenine ring had a profound effect on the calculated isotope effects for both functionals. Also between the two functionals, differences in the calculated KIEs were found. Generally M06-2X suggested larger isotope effects then B3LYP for the different models. Discrepancies between functionals and basis sets used to generate the models has been reported previously by Singh and Schramm for *Streptococcus pneumonia* 5'-methylthioadenosine nucleosidase (MTAN) where a variety of methods (ab initio, HF, and DFT), and basis sets (3-21G, 6-31G(d), and 6-31G(d,p)) were used to calculate KIEs of the transition state.[124] The 1'-^{14}C primary isotope effect was predicted well by most method and basis set combinations; however, the predicted 9-^{15}N primary nitrogen isotope effect ranged from 1.035 (HF/3-21G and B1LYP/6-31G(d)) to 1.024 (HF/6-31G(d,p)).

Initially the PRATP structure was truncated to include only the adenine ring of ATP, with the ribose part replaced by a hydrogen. This was done under the assumption that atoms more than two bonds away would not affect the KIE. However, the results from the different charge and tautomer states suggest that electronic effects further away do affect the result through the conjugated ring system. The N9 proton was thus replaced by the ribose ring

part in both the ground state and the dissociated adenine base. The lowest energy ground state was again found by starting from several conformations, generated by rotating the *N*-ribosidic bond. This gave back a single resultant conformation for all structures tested. In the dissociated base, the dihedral angle to the ribose ring was fixed to represent the angle observed in the ATP bound crystal structure of *M. tuberculosis* ATP-PRTase (Section 2.8.3). The different charge and tautomeric states were generated and the structures optimised at the B3LYP/6-31^{+}G(d,p) level (Figure 4.6). The B3LYP functional was chosen based on results discussed in Section 4.5.3. The corresponding calculated KIE results can be found in Table 4.3.

Table 4.3. The calculated KIEs for a dissociated adenine base including the ATP Ribose and the effects of charge and tautomer state on the kinetic isotope effect. Experimental KIEs included for comparison.

	B3LYP			Experimental KIE	
State	**1-^{15}N**	**6-^{15}N**	**ATP-PRTase**	**1-^{15}N**	**6-^{15}N**
Neutral	1.018	0.992	***M. tuberculosis***	1.024 ± 0.001	1.012 ± 0.001
Anionic	1.024	1.005			
Neutral N3 tautomer	1.019	1.000	***C. jejuni***	0.995 ± 0.003	1.005 ± 0.001
Neutral N7 tautomer	1.023	1.007			
Protonated at N3	1.013	0.981	***L. lactis***	1.028 ± 0.001	0.996 ± 0.003
Protonated at N7	1.017	0.988			

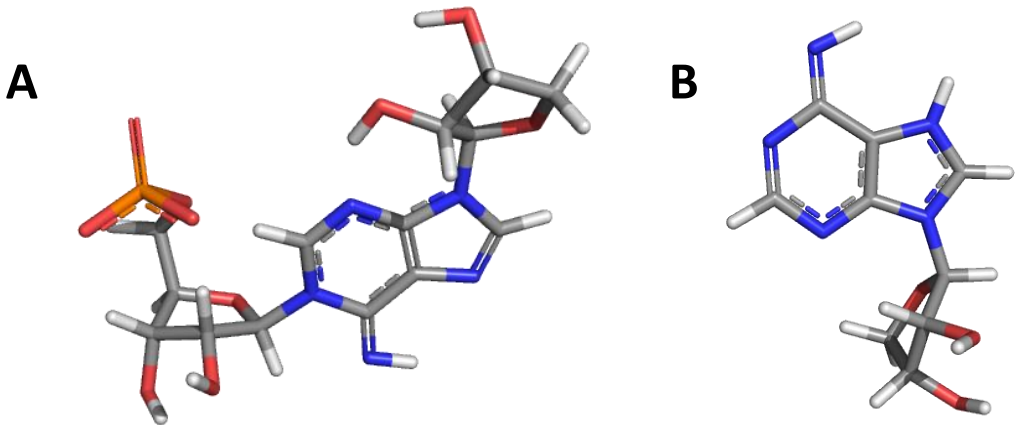


*Figure 4.6. Adenine with ribose expansion. **A** The extended ground state model. **B** The extended neutral N7 tautomer as an example of the extended adenine base.*

Inclusion of the ribose ring did not give rise to large changes in the calculated nitrogen KIEs. This suggests that the simplified model was sufficient to capture the features of the transition state of the dissociation of the adenine ring and the results in Table 4.2 will therefore be used to discuss the nitrogen isotope effect.

Matching the experimental and calculated KIEs indicates that an anionic or neutral (as one of its tautomers) were possible states for the adenine ring at the transition state of *M. tuberculosis* and *L. lactis* ATP-PRTase. The formation of altered tautomers or an anionic adenine requires residues which can protonate or deprotonate the adenine, especially considering the physiological direction (PRPP + ATP → PRATP + PP_i). Crystal structures with ATP bound had been obtained for *M. tuberculosis* ATP-PRTase and *C. jejuni* ATP-PRTase (discussed in Chapter 2) and were examined for such residues.

M. tuberculosis ATP-PRTase has a hydrophobic pocket where the adenine part binds. A hydrogen bond is made to a water molecule at the N3 position; this same water molecule also coordinates to Mg^{2+} (Figure 4.7, **A**). No further polar interactions occur and no other active site residues that could aid in deprotonation of the adenine or stabilise the anion are in close proximity.

A slightly different binding mode and different interactions were observed for the ATP bound *C. jejuni* ATP-PRTase structure. The adenine is further stabilised by Arg16 which forms a sandwich π-π stacking interaction enclosing the adenine ring between Leu170 and Arg16. In addition the adenine amine binds in close proximity to Ser172 (Figure 4.7, **B**).

The ATP binding site of *L. lactis* ATP-PRTase contains features from both *C. jejuni* and *M. tuberculosis* ATP-PRTase. An arginine is situated close to the adenine binding envelope, similar to Arg16 in the *C. jejuni* ATP-PRTase structure; however this does not greatly alter the binding position of the adenine part of ATP between the two structures. The triphosphate binding site is structurally more similar to *M. tuberculosis* ATP-PRTase and gives rise to a different binding mode around the triphosphate part. The *M. tuberculosis* ATP-PRTase crystal structure with ATP bound was thus used to generate an overlay of the catalytic core (domain I and II, residues 1–200) of the *M. tuberculosis* structure and the active site of *L. lactis* ATP-PRTase (residues 1–204). This resulted in an RMSD of 1.601 Å over 141 atoms. (Figure 4.7, **C** and **D**). The overlay revealed that in addition to a hydrogen bond

to water at N3 and a π-π stacking interaction from Arg10, Glu158 was in close proximity to both the N6 and N7 of the adenine ring and therefore potentially able to interact at those positions.

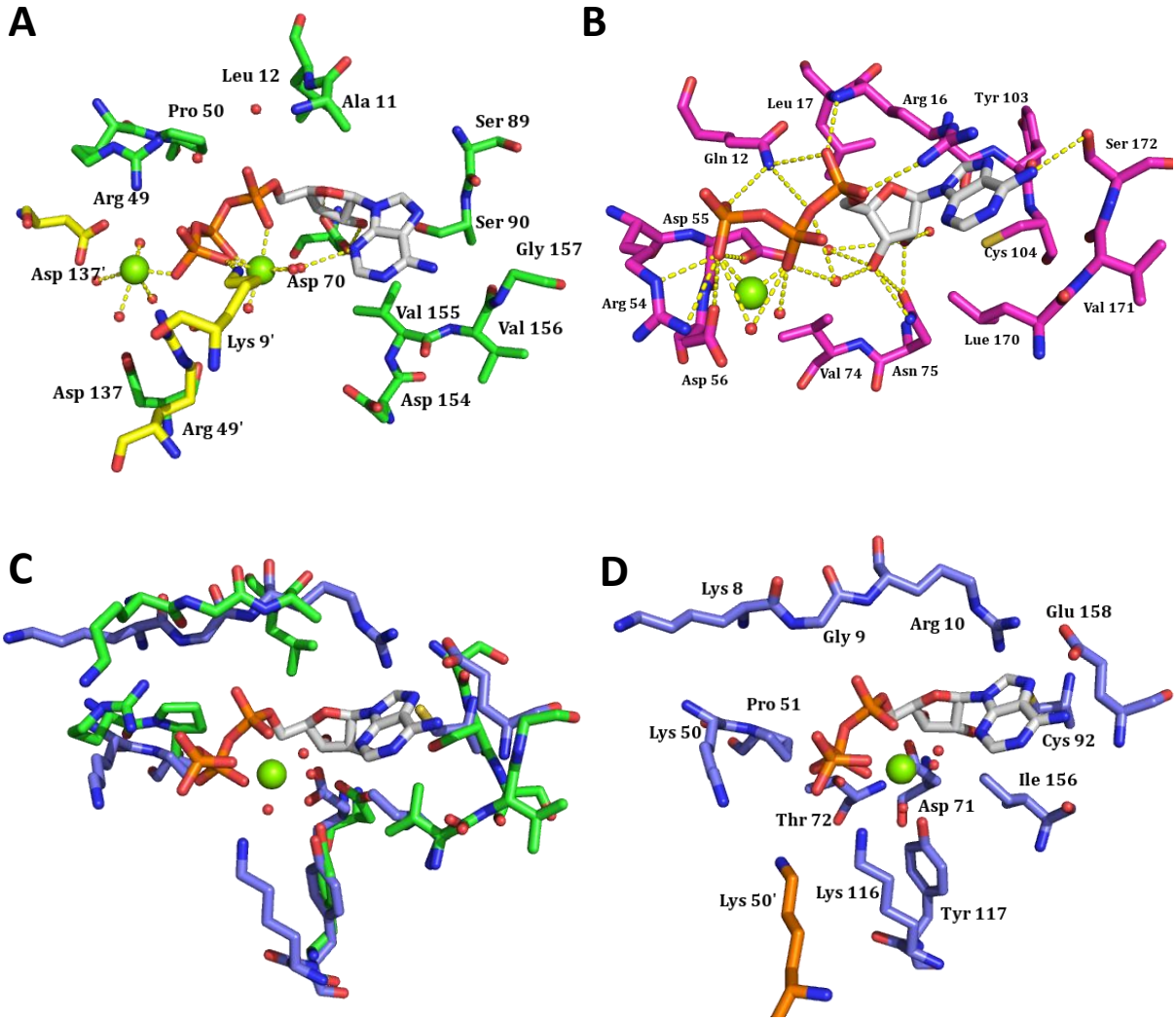


Figure 4.7. ATP binding. ***A*** *Active site of M. tuberculosis ATP-PRTase (main chain green, second chain yellow) with ATP bound. H-bonds as yellow dashed lines.* ***B*** *Active site C. jejuni ATP-PRTase (magenta) with ATP bound at the active site.* ***C*** *Active site alignment of M. tuberculosis ATP-PRTase (green) with ATP bound and L. lactis ATP-PRTase (blue) (PDB: 1Z7N, PRPP not shown).* ***D*** *The L. lactis ATP-PRTase active site (main chain blue, adjacent active site orange) with ATP displayed from the alignment with M. tuberculosis ATP-PRTase.*

The nitrogen isotope effects for *M. tuberculosis* were 1.024 ± 0.001 for 1-^{15}N and 1.012 ± 0.001 for 6-^{15}N (Section 3.3.3.1). The anionic and neutral N3 and N7 tautomers agree best with these values, with large primary nitrogen isotope effects and smaller positive 6-^{15}N isotope effects. Based on the crystal structure the anionic adenine was considered an

unlikely state, with no residues in close proximity able to deprotonate the adenine. Furthermore, stabilisation through a hydrogen bonding network was ruled out, with few of these interactions made to the adenine part. Of the neutral tautomeric states the N3 tautomer was considered the most likely state based on the close proximity of a resolved water molecule in the crystal structure of *M. tuberculosis* ATP-PRTase, held in place by Mg^{2+}. The detail in the crystal structure was not considered sufficient to distinguish between the E or Z conformationof the imine proton. The small discrepancy between the calculated and measured isotope effect could arise from factors such as the used functional and/or basis set, or the unmodelled active site interactions. Although the N3 tautomer was the more likely state the N7 tautomer cannot be ruled out based on these results. Overall the results support a fully dissociated adenine at the transition state for all charged and tautomeric states and the transition state was modelled accordingly.

L. lactis ATP-PRTase gave a slightly larger primary 1-^{15}N isotope effect of 1.028 ± 0.001 and suggests a substantial increase in vibrational freedom following dissociation of the *N*-glycosidic bond (Section 3.3.3.3). The 6-^{15}N isotope effect near unity at 0.996 ± 0.003 suggests no significant change in vibrational freedom upon dissociation. The best match between calculated and experimental KIE was for an anionic adenine or N7 tautomer. At the active site of *L. lactis* ATP-PRTase Arg10 is likely to interact with the adenine through sandwich π-π stacking, similar to what is observed for *C. jejuni* ATP-PRTase. These types of interactions have been shown to play a role lowering the pK_a tyrosine as well as purine bases.[129,130] Glu158 is part of the phosphate binding loop, and is in close proximity to the adenine base. It is therefore potentially able to interact and stabilise an anionic ring system or alternate tautomeric state. However, given the pK_a of ~16 for the adenine amine, the anionic adenine was considered less likely to occur even if multiple interactions would facilitate this step.[131,132] Overall, the small differences between the model and experimental KIE could arise from the discussed un-modelled interactions or the functional and/or basis set used. The results support a fully dissociated adenine base at the transition state of *L. lactis* ATP-PRTase.[85]

C. jejuni ATP-PRTase gave a primary isotope effect near unity of 0.995 ± 0.003 (Section 3.3.3.2). An isotope effect near unity suggests there is no change in vibrational freedom at the transition state through bond breaking and/or that several bonding changes fortuitously

compensate for each other in the conjugated purine ring.[85,104] Given the similarities of the active site of *L. lactis* and *C. jejuni* ATP-PRTase around the adenine binding pocket, this result is surprising and cannot be explained through compensation by several bonding changes alone. Another explanation would be that *N*-glycosidic bond cleavage does not occur close to the rate determining step and is thus suppressed. Given the similarities of the primary ^{14}C and α-secondary ^{3}H isotope effects to the other two ATP-PRTases studied, a fully dissociated adenine ring was also assumed for this enzyme.

4.5.2 The transition state model generation

The Hammond postulate suggests that the transition state resembles the species closest in free energy.[133] With the nitrogen isotope effect suggesting a dissociated nitrogenous base the transition state was expected to resemble the ribose 5-phosphate-phosphonoacetic acid (PRPA) species.

It is important to consider during model generation that the enzyme active site can introduce conformational constraints to the substrate or product during catalysis. A crystal structure can provide insight into such constraints as well as any other factors playing a role during catalysis such as metals, active site residues, and conserved water molecules. However, no crystal structure with PRPA bound at the active was available. There is a PRPP bound crystal structure of ATP-PRTase from *L. lactis* available, which is expected to closely resemble the PRPA species, although at lower resolution (3.0 Å PDB:1Z7N). Low resolution can make accurate modelling of ligands challenging and information from such structures was used with caution. The ligand binding modes across the four chains were compared, with differences in geometries observed for the four ligands in particular around the pyrophosphate part (Figure 4.8). The omit map (Fo-Fc) further emphasises the poor fit of these ligands and the PRPP conformation from these structure were therefore considered a bad starting point for the transition state structure.

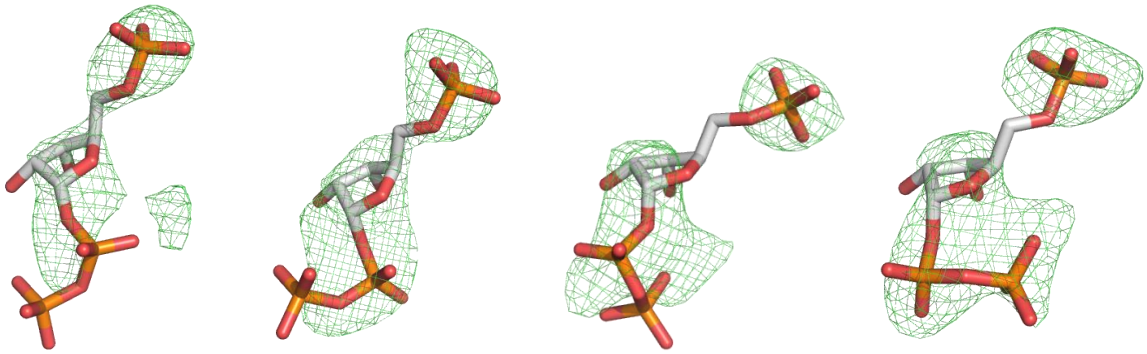


Figure 4.8. The different binding modes for PRPP modelled at the active site of L. Lactis. The omit map (Fo-Fc) is displaced in grey (1Z7N).

Type I and Type IV PRTase enzymes share a signature fold responsible for PRPP binding (Section 1.10). Because of this high level of conservation a similar binding mode for PRPP was expected. To test this hypothesis the PDB database was searched for crystal structures of Type I phosphoribosyl transferase enzymes with PRPP bound. Structures with resolutions greater than 2.50 Å were excluded from this search. A total of 10 structures meeting these criteria were found and an overlay of the phosphate binding loop, which is part of this signature fold, was generated (Figure 4.9). From the overlay it can be seen that PRPP adopts a very similar binding mode across the various species examined. In addition, for most structures the catalytically essential Mg^{2+} cation was present situated under the ribose ring coordinated to the pyrophosphate, the 2- and 3-OH, and water. Magnesium is required for activity and the observed binding mode make it likely to play an important role at the transition state. The metal ion was therefore also included in the model. The binding mode of PRPP bound to adenine phosphoribosyltransferase from *Giardia lamblia* (PDB: 1L1R) including magnesium was used as the starting point for the transition state model.

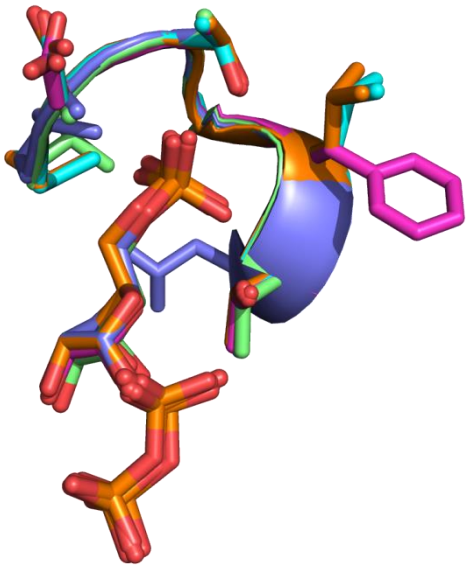


Figure 4.9. Overlay of the phosphate binding loop of some Type I PRTase enzymes with the PRPP product bound. PDB: 1TC2, 1FSG, 1L1R, 1P18 and 1ZN7 shown.

The pyrophosphate group was replaced by the alternative substrate phosphonoacetic acid and the structure optimised without constraints using B3LYP and the 6-31G(d) basis set. This model was unstable with the ligands around the magnesium reorganising, the 5-phosphate formed unrealistic interactions to other parts of the model and in some cases the model did not converge (Figure 4.10, **A**). A crystal structure has been obtained for human hypoxanthine-guanine phosphoribosyl-transferase (HGPRTase) with a transition state analogue bound at the active site including pyrophosphate.[134] The pyrophosphate is bound in the same orientation around magnesium as observed for PRPP. The observation of the ligands rearranging around the magnesium in the transition state model were thus considered unrealistic.

To reduce unrealistic interactions the ribose was truncated by removing the ribose 5-phosphate part. Optimisation of the ribose ring and divalent magnesium cation separately prevented the rearrangement. After the ligands around the magnesium were optimised the ribose 5-phosphate could be reincluded.

This iteration was an improvement from the first model, however when the C1-O^{PA} bond was further extended this still resulted in unrealistic and unstable models. More specifically, interactions of the phosphate with ligands around the magnesium would result in unrealistic rearrangements around the magnesium ion (Figure 4.10, **B**, **C** and **D**) or the abstraction of the 2-OH proton (Figure 4.10, **C**).

At the active site the charged ribose 5-phosphate is accommodated by the phosphate binding loop, anchoring PRPP in place for catalysis.[38] To resemble these constraints imposed by the active site, additional constraints by fixing the dihedral angle to the ribose 5-phosphate, have been reported.[61] The dihedral angles to this group were fixed and the phosphate protonated to resemble the charged interaction formed upon binding at the active site. The structure optimisation calculation was then repeated. This gave a stable transition state model that could be extended without becoming unstable. The absence of any additional imaginary frequencies suggests that fixing the dihedral angle to the 5-position did not introduce substantial unrealistic constraints to the system.

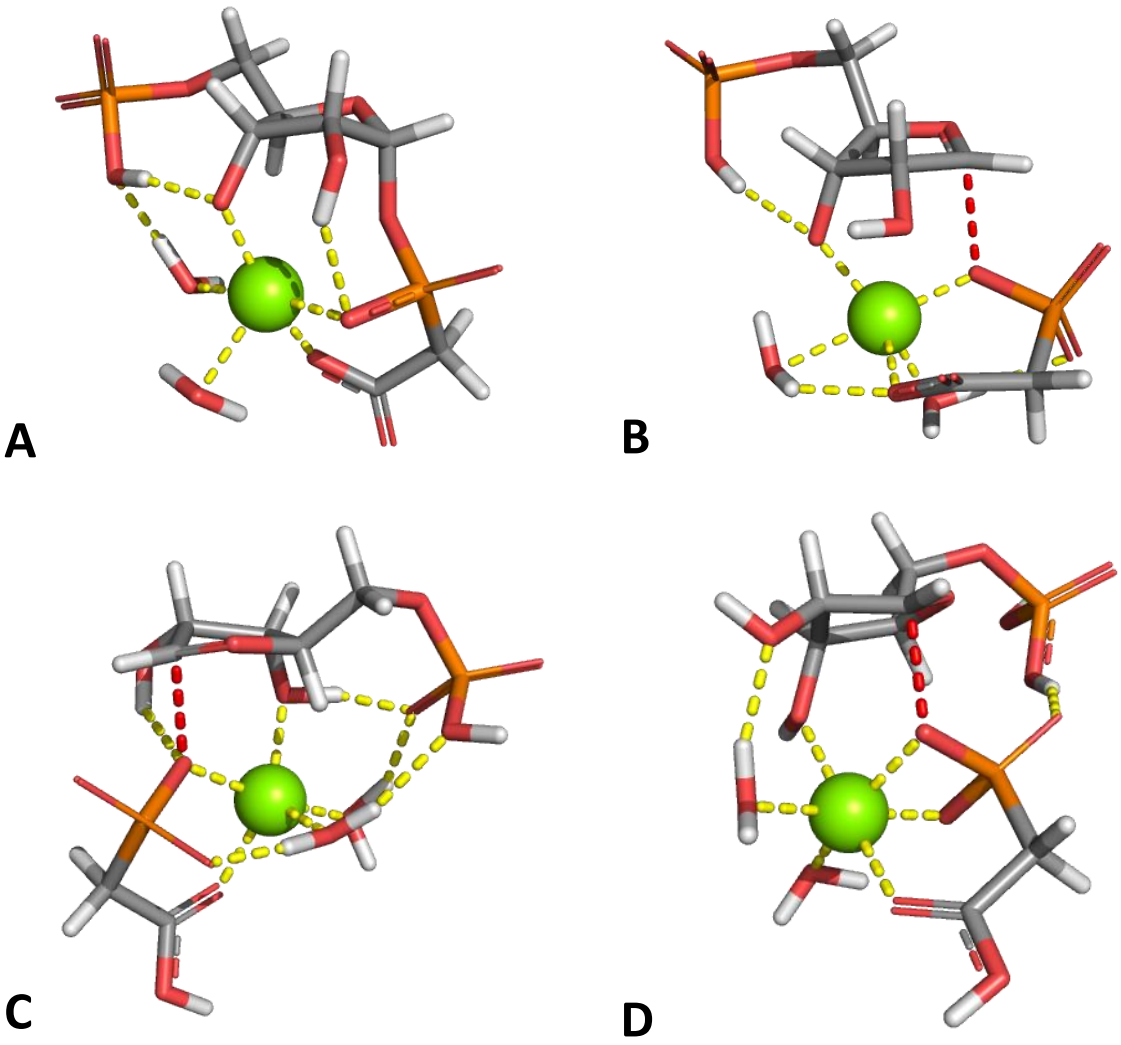


Figure 4.10. PRPA after the 5-phosphate is reincluded. Yellow dotted lines indicate interactions between atoms, the red dashed line shows the scissile bond. ***A*** *The PRPA model after optimisation with a dianionic PA and deprotonated phosphate input structure.* ***B*** *Deprotonation by the phosphate group and magnesium has lost the more favourable octahedral configuration.* ***C*** *PRPA partially protonated, Mg has lost the octahedral coordination configuration.* ***D*** *PRPA fully protonated, the ligands around magnesium have rearranged upon elongation of the C1'-O^{PA} bond.*

4.5.3 Different functionals for the transition state structure calculations

In the calculations of the nitrogen KIEs differences where seen between the B3LYP and M06-2X functionals. A similar affect was expected for the optimisation of the transition state with these two functionals. To establish how well the two functionals could model the change in bonding interactions, the C1'-O^{PA} bond (Figure 4.11) was iteratively extended by 0.1 Å until loss of the bond was observed (no or small imaginary frequency). A gradual change in the force constants (Section 3.1.4), in particular for C1', are expected upon extension of the C1'-O^{PA} bond and the sigma bond is broken (Figure 4.12).

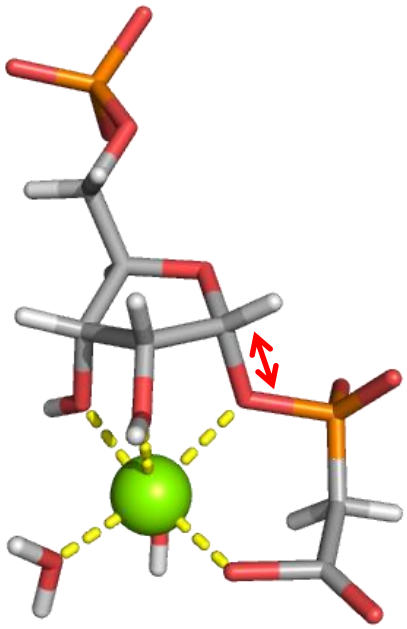


Figure 4.11. PRPA input model. the red arrow indicates the C1'-O^{PA} bond was iteratively extended.

B3LYP and M06-2X gave a similar response for the α-secondary 1'-^{3}H KIE with B3LYP predicting a more gradual increase in conformational freedom then M06-2X. A stronger difference was predicted for the primary 1'-^{14}C KIE. The M06-2X functional showed little change in the primary KIE for bond lengths of up to 2.10 Å after which a rapid change was predicted with full loss of interactions past 2.3 Å. A bond distance of 2.1 Å corresponded to a natural bond order of ~0.33 or loss of over 67% of the original bonding interactions. It was considered unlikely for this to not have an effect on the primary isotope effect. Furthermore the full loss of bonding interaction over 0.2 Å is unrealistic. A possible explanation for the underperformance of the M06-2X functional compared to B3LYP is that it was

parameterised for non-metals.[107,118] Magnesium has a central role in the transition state model and potentially the incorrect modelling of the electron density around the magnesium may have caused the observed effect.

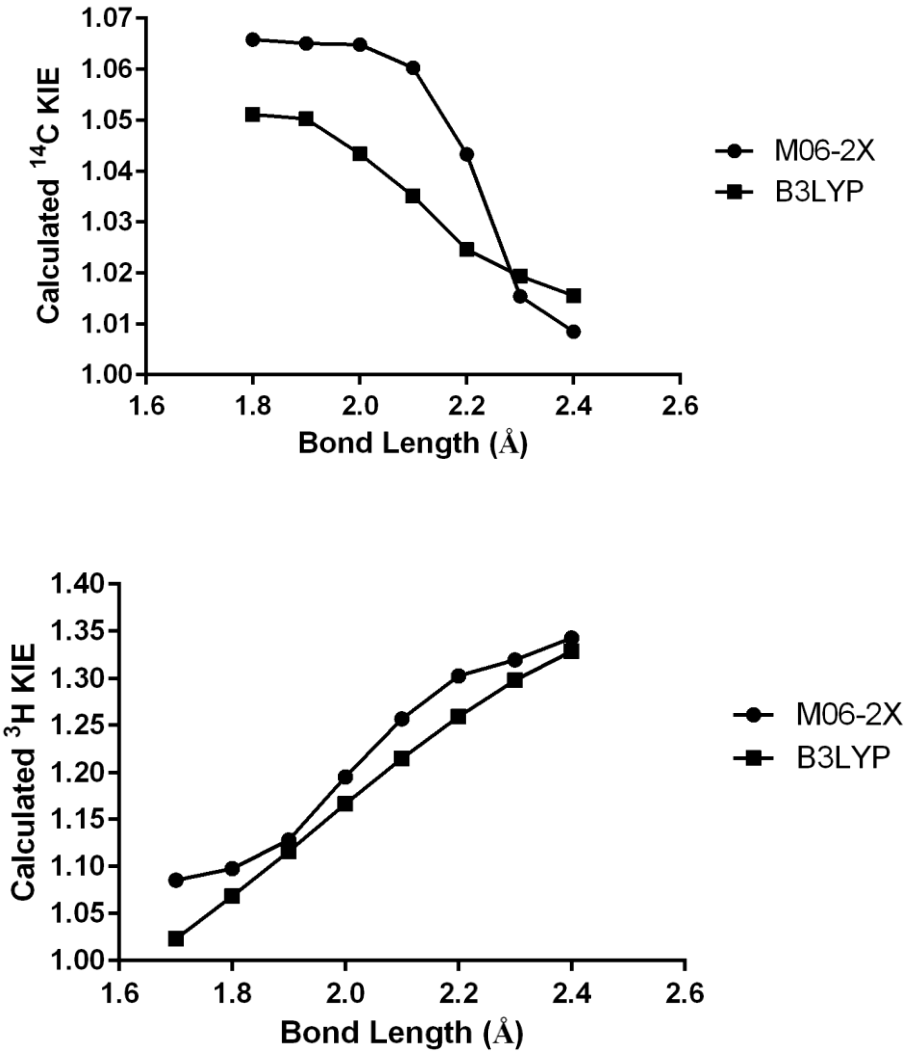


Figure 4.12. Relationship between bond length (Å) and calculated KIE for the primary ^{14}C and the α-secondary ^{3}H isotope effect for the two functionals tested.

B3LYP was considered to perform well; but starts to fail in some cases to predict long range interactions. To test if an improvement on the B3LYP functional could be achieved by including a long range correction, the CAM-B3LYP functional was also included. The CAM-B3LYP combines the hybrid quality of B3LYP and the long range correction presented by Tawada *et al.*, who suggested the treatment of short range exchange interaction by means of an appropriate pure density functional and the long-range interaction by Hartree-Fock

(HF) exchange.[135,136] To explore the trend in the calculated KIE the TS model calculations were repeated with the CAM-B3LYP functional and the 6-31G(d) basis set (Figure 4.13).

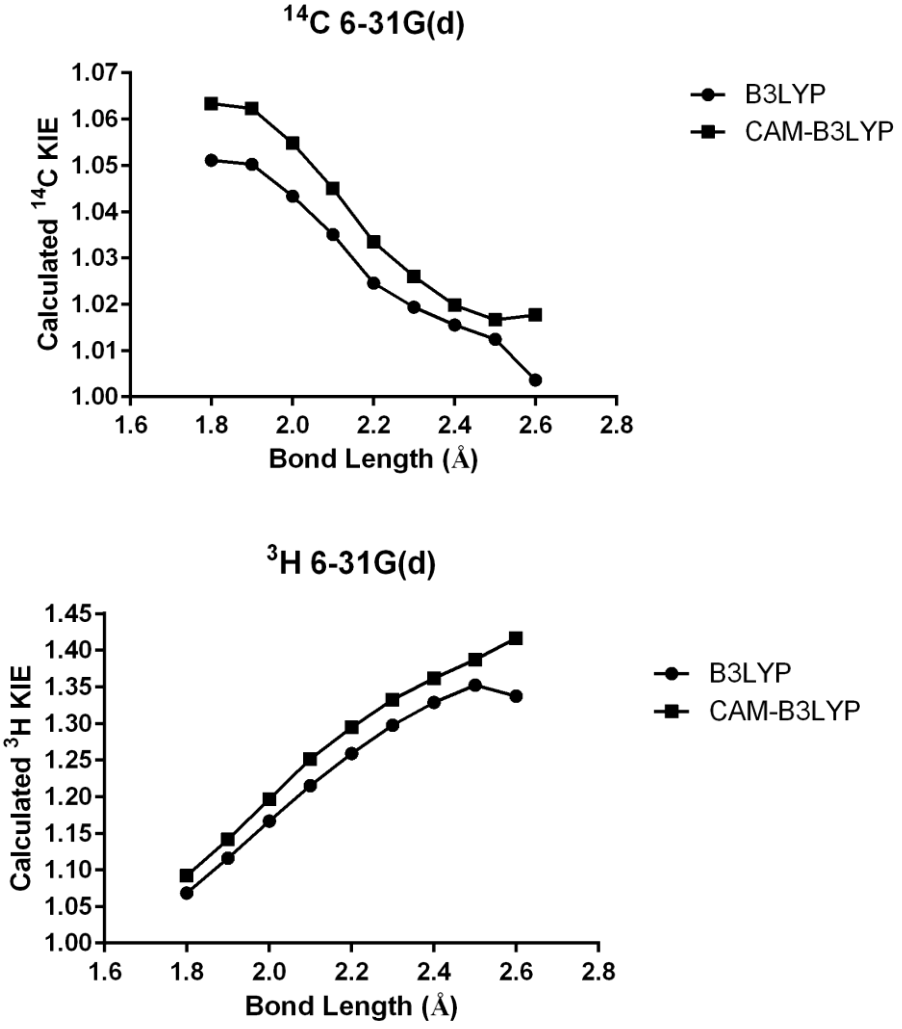


Figure 4.13. Results of the grid search with B3LYP and CAM-B3LYP functionals with the 6-31G(d) basis set for both ^{14}C and ^{3}H.

The predicted primary ^{14}C KIE at shorter bond lengths was increased relative to the isotope effects predicted for B3LYP reflecting on this difference in functionals. The longer range interaction up to 2.5 Å gave a similar performance, with a gradual loss of the predicted primary ^{14}C kinetic isotope effect upon the formation of a greater carbocation being characteristic for both functionals. Beyond this bond length the B3LYP functional started to fail with complete loss of the interaction between the C1'-O^{PA} bond. The CAM-B3LYP functional gave a slight increase in the primary isotope effect but still modelled this interaction reasonably well based on the observed trend. Overall both functionals predicted primary isotope effects higher than expected for near full bond loss, with primary isotope

effects expected near unity for a carbocation intermediate.[124] The limited basis set may have been responsible for this result. To increase modelling accuracy for the atomic orbitals more distant from the nuclei a diffuse function can be added to the basis set. The basis set was thus extended to include a diffuse function on the heavy atoms. A polarization function was also added to the hydrogen atoms in order to give flexibility to the s-orbital. The ground state model was also optimised with the extended basis set to ensure appropriate comparison with the ISOEFF98 program. The increase in basis set had a similar effect on both functionals with isotope effects predicted closer to the expected value of near unity (Figure 4.14). The α-secondary followed a similar trend in both functionals with similar isotope effects predicted across the bond lengths investigated.

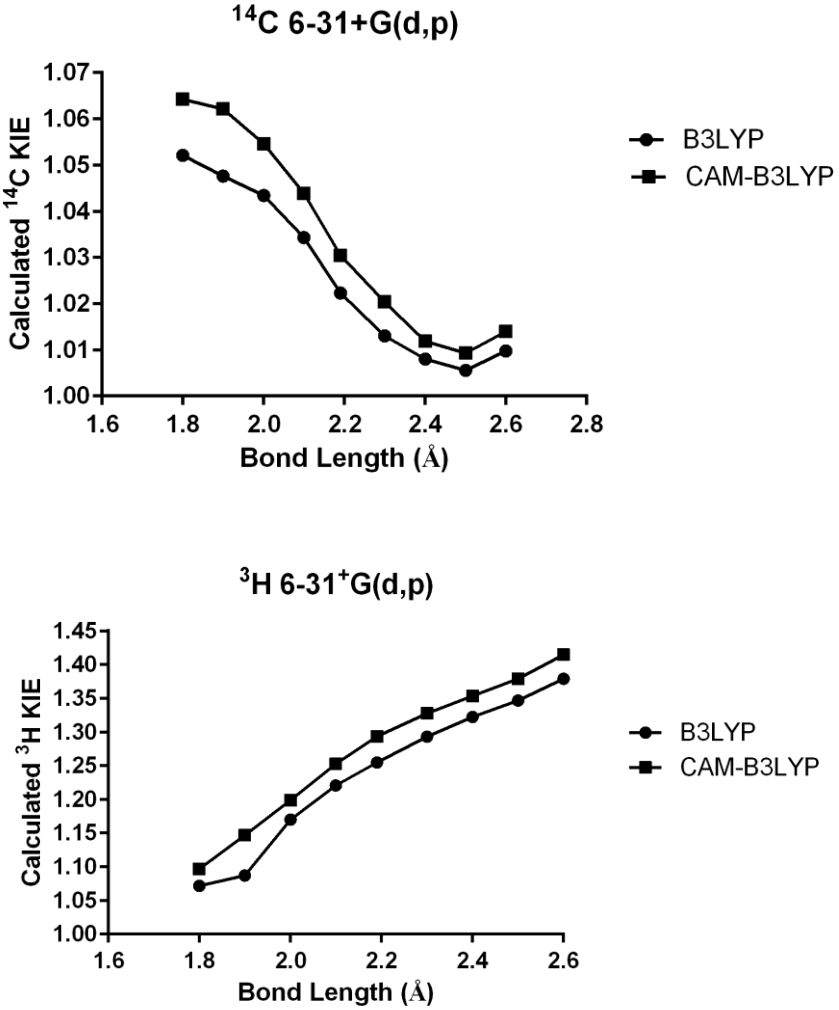


Figure 4.14. Results of the grid search with B3LYP and CAM-B3LYP functionals and the extended 6-31^{+}G(d,p) basis set for both ^{14}C and ^{3}H.

Overall the CAM-B3LYP did not give significant improvement over the B3LYP functional, in particular around the area were the transition states for ATP-PRTase enzymes were expected (1′-^{14}C intrinsic KIEs of 1.031–1.028). In addition B3LYP has the most precedent in literature for this type of reaction mechanism and was thus the preferred functional for further refinement of the transition state models.[61,85,126]

Phosphonacetic acid was fully protonated during the transition state model generation to prevent unrealistic interactions upon extension of the *N*-ribosidic bond. Phosphonacetic acid has a pK_1 = 2.0, pK_2 = 5.1, and pK_3 = 8.7. At pH = 8.5 the carboxylic acid was therefore expected to be deprotonated.[137] From Figure 4.14 a transition state with a C1′-O^{PA} of >2.1 was expected. Starting at this bond length the transition state structure was further refined with an anionic PA nucleophile. Deprotonation of the carboxylic acid did not have any adverse effect on the transition state model, but did significantly affect the calculated primary isotope effect. The anionic PA was expected to be the most realistic representation of the transition state for the ATP-PRTase enzymes and thus used to find a good match to the experimental KIE values.

After refinement the best match for *M. tuberculosis* was at a C1′-O^{PA} bond length of 2.75 Å. The bond length for *C. jejuni* ATP-PRTase was predicted to be close, at 2.65 Å. Also the short form of the ATP-PRTase enzyme from *L.lactis* gave a very similar late stage transition state with a C1′-O^{PA} bond length of 2.50 Å (Table 4.4) (Appendix . The substrates in all three enzymes show carbocation characteristics with a positive charge build up at the anomeric centre which is stabilised by the ribose ring oxygen, highlighted by the Mulliken charges (Table 4.5). Mulliken charges provide a means of estimating partial atomic charges from QM calculations.[138,139]

Table 4.4. Results from the refined search for the transition state structures of M. tuberculosis, C. jejuni and L. lactis ATP-PRTase.

ATP-PRTase from	C-O bond distance	1′-^{3}H Calculated	1′-^{14}C Calculated	1′-^{3}H Experimental	1′-^{14}C Experimental
L. lactis	2.50 Å	1.407	1.031	1.147 ± 0.008	1.031 ± 0.001
C. jejuni	2.65 Å	1.441	1.029	1.212 ± 0.002	1.029 ± 0.001
M. tuberculosis	2.75 Å	1.457	1.028	1.250 ± 0.008	1.028 ± 0.002

Table 4.5. The predicted Mulliken charges of the primary and α-secondary position to support the changes going from the ground state to the transition state. The ring oxygen was included to show the change in charge at this centre.

ATP-PRTase	1′-C	1′-H	Ribose ring O
L. lactis	0.253	0.226	-0.013
C. jejuni	0.275	0.235	0.004
M. tuberculosis	0.290	0.236	0.019
GS	-0.560	0.112	-0.302

4.5.4 Discussion of the transition state model

4.5.4.1 Transition state model for *Mycobacterium tuberculosis* ATP-PRTase

The transition state was matched closely to the primary 1′-^{14}C intrinsic kinetic isotope effect, because not only does this position report most accurately on the reaction mechanism, it is also insensitive to geometry changes upon binding and is expected not to be affected by interactions at the active site.[140] It was found that a C1′-O^{PA} bond distance of 2.75 Å was in good agreement with the primary isotope effects. The calculated α-secondary KIE was substantially larger than the experimental value. This sometimes arises from over estimation by in vacuo computation of the KIE at this position due to uncertainty around the H1′ position.[61,102,104,124] A bond length of 2.75 Å corresponds to a late associative D_N*$A_N^{\ddagger}$ transition state (Section 4.2 and Figure 4.15). Going from the ground state to the transition state there is a positive charge build-up at the C1′ position as indicated by the change in Mulliken charge of -0.560 compared to 0.290 at the transition state. The ribose ring oxygen stabilises the positive charge, resulting in the generation of a double bond characteristic, as this centre undergoes a hybridisation change from sp^3 towards sp^2, characteristic for a carbocation species. The Mulliken charge again highlights this change with reduced negativity at the oxygen atom. The double bond characteristic is also reflected in the C1′-O^{ribose} bond length which is shortened by 0.18 Å from 1.43 Å in the ground state to 1.25 Å at the transition state. The divalent magnesium cation plays an important role in placing the incoming nucleophile in the right position by coordinating the nucleophilic phosphonate oxygen and carboxylic acid functionality. When the MEPS map of the transition state is compared to the ground state model, the difference in charge around the reaction centre is

emphasised with an electron deficient anomeric carbon and also a less electron negative ring oxygen (Section 4.4 and Figure 4.15) (Section 8.4.5).

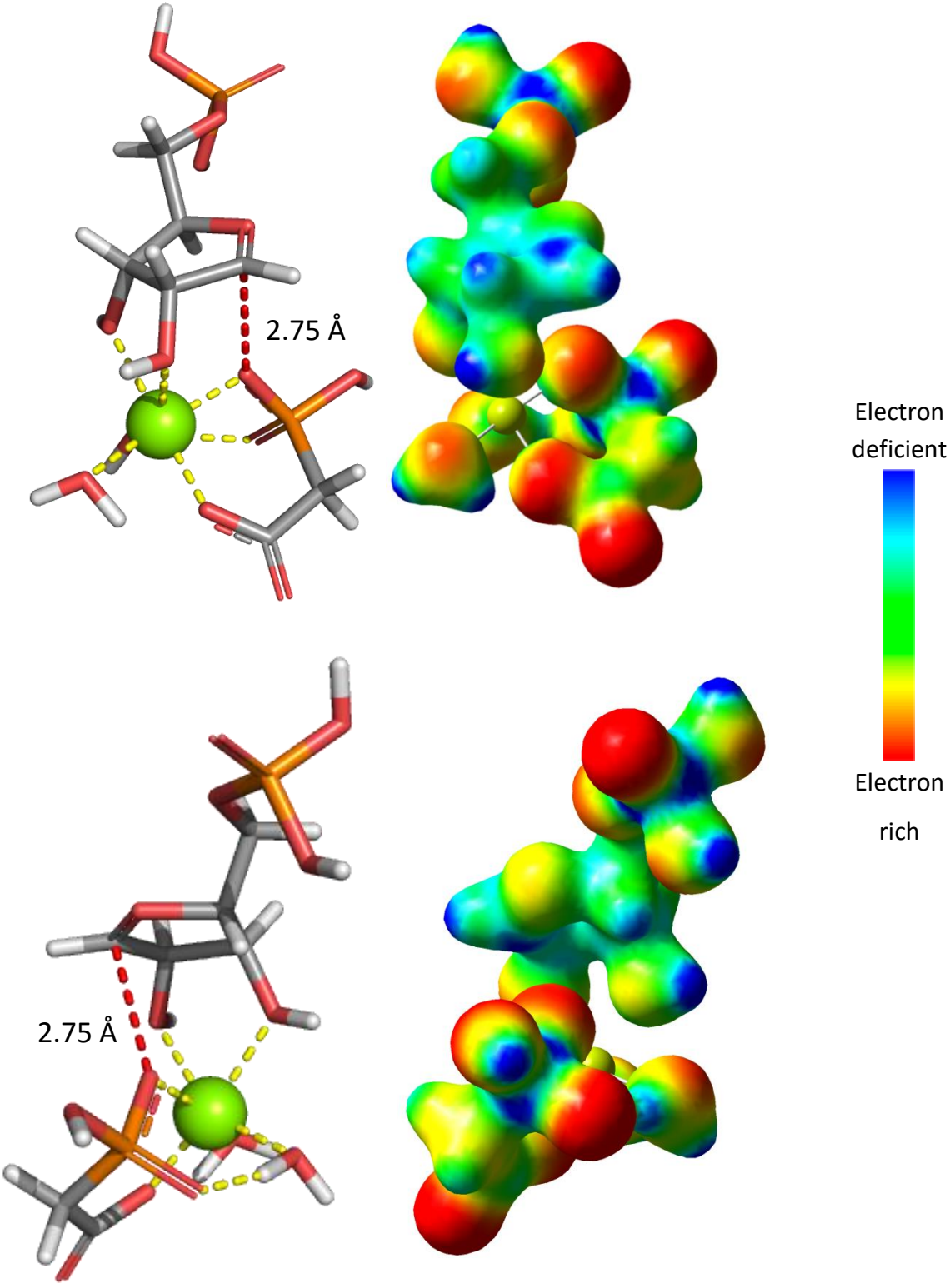


Figure 4.15. The transition state structure of M. tuberculosis ATP-PRTase and the corresponding MEPS map, in two orientations. Scale gives the colour coding with electron rich parts indicated in red and electron deficient in blue.

4.5.4.2 Transition state model for *Campylobacter jejuni* ATP-PRTase

The best match of the calculated KIEs to the intrinsic primary ^{14}C KIE was found for a transition state structure with a C1'-O^{PA} bond length of 2.65 Å. For this transition state model the calculated α-secondary KIE was substantially larger than the experimental value attributed to condensed phase uncertainty around the C1'-H1' region.[61,102,104,124] The anomeric carbon at this bond length has become more positive compared to the ground state, with an increase in Mulliken charge of 0.85 (Table 4.5). This change also reflects on the α-secondary proton with an increase in positive charge. The ribose ring oxygen stabilises the positive charge at the anomeric carbon by sharing its lone pair of electrons, shortening the C_{α}-O_{ribose} bond from 1.43 Å to 1.26 Å. The double bond characteristics between the anomeric carbon and ribose ring oxygen has a ring flattening effect. The carbocation is also evident from the electrostatic potential map with a more electron deficient anomeric carbon. The sharing of the lone pair of electrons has reduced the electron richness of the ring oxygen compared to the ground state structure (Section 4.4 and Figure 4.16). The octahedral coordinated magnesium plays a vital role in placing the phosphonate under the anomeric carbon. Overall the transition state model supports a late associative $D_N*A_N^{\ddagger}$ transition state for the ATP-PRTase enzyme from *C. jejuni* (Section 8.4.6).

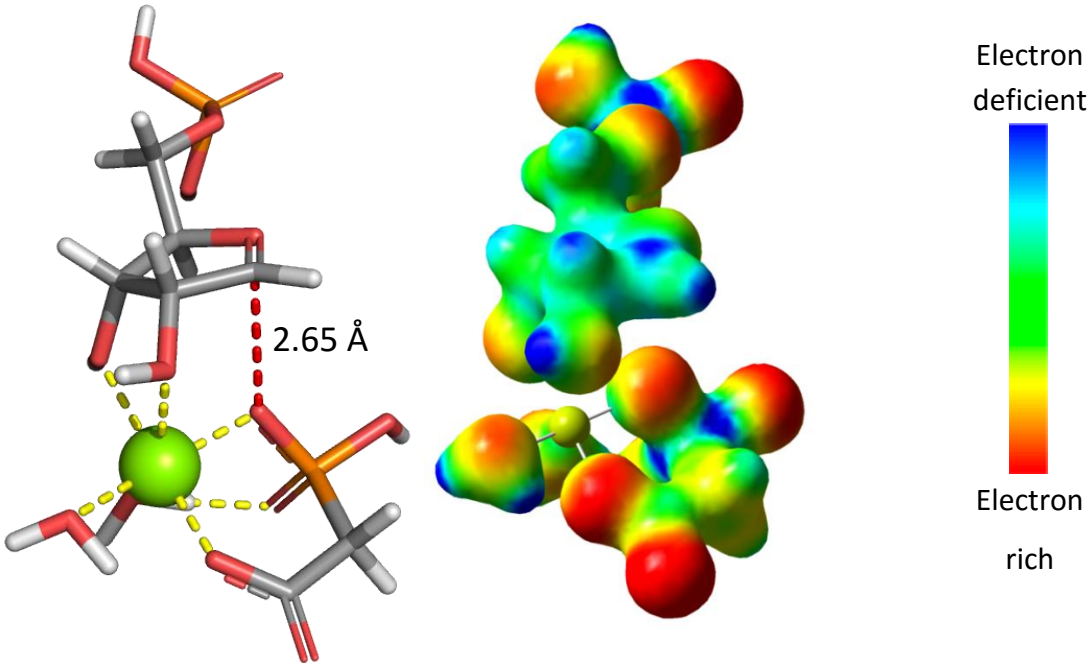


Figure 4.16. The transition state structure of C. jejuni ATP-PRTase and the corresponding MEPS map, in two orientations. Scale gives the colour coding with electron rich parts indicated in red and electron deficient in blue.

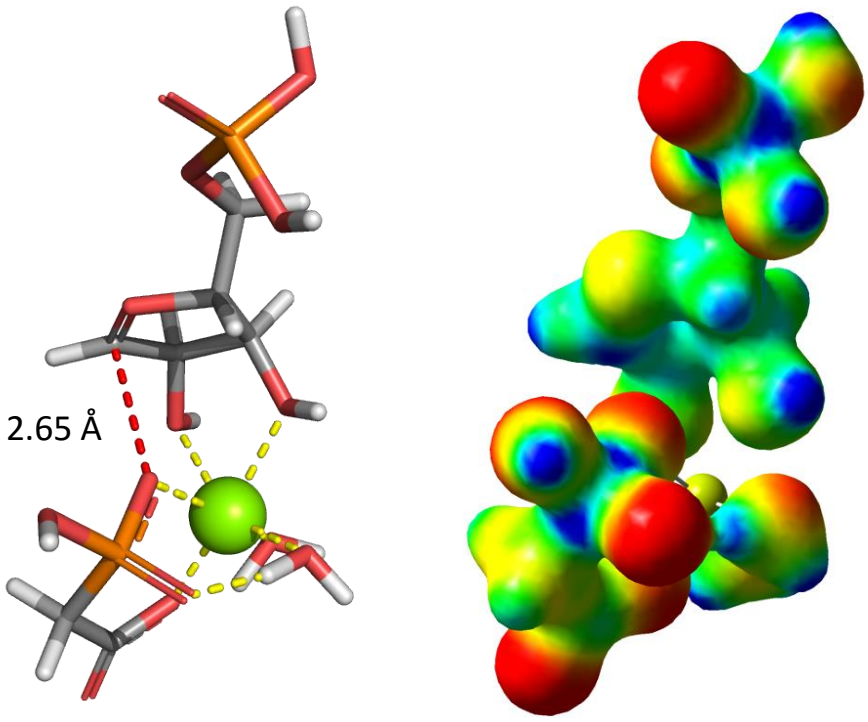


Figure 4.16. *(continued)*

4.5.4.3 Transition state model for *Lactococcus lactis* ATP-PRTase

The short form ATP-PRTase enzyme from the organism *L. lactis* was found to go through a very similar transition state to the long form ATP-PRTase enzymes discussed above. Here, the best match to the intrinsic isotope effects was for a transition state with a C1'-O^{PA} bond length of 2.50 Å. This corresponds to a late associative $D_N{*}A_N^{\ddagger}$ transition state (Section 4.2). The calculated α-secondary KIE was slightly larger than the experimental value due to overestimation of the KIE at this position by *in vacuo* computation, due to condensed-phase uncertainty around the C1'-H1' region.[61,102,104,124] Furthermore, constraints from van der Waals interactions at the active sites are proposed to suppress the experimental 1'-3H KIEs.[61] The Mulliken charge change at the anomeric carbon from -0.560 to + 0.253 illustrates the carbocation characteristics. Participation by lone pair electrons of the ribose ring oxygen helps stabilise the positive charge of the carbocation species (Table 4.5). This leads to a double bond characteristic, with shortening of the bond from 1.43 Å in the ground state to 1.26 Å at the transition state, and flattening of the ribose ring. The observed changes at this centre also reflect on the α-secondary proton with an increase in Mulliken charge from 0.112 to 0.226. The MEPS map shows the electron deficient anomeric carbon and the

anomeric proton (Figure 4.17). The electrostatic potential around the ring oxygen has changed from an electron rich position in the ground state to a more neutral position at the transition state (Section 8.4.7).

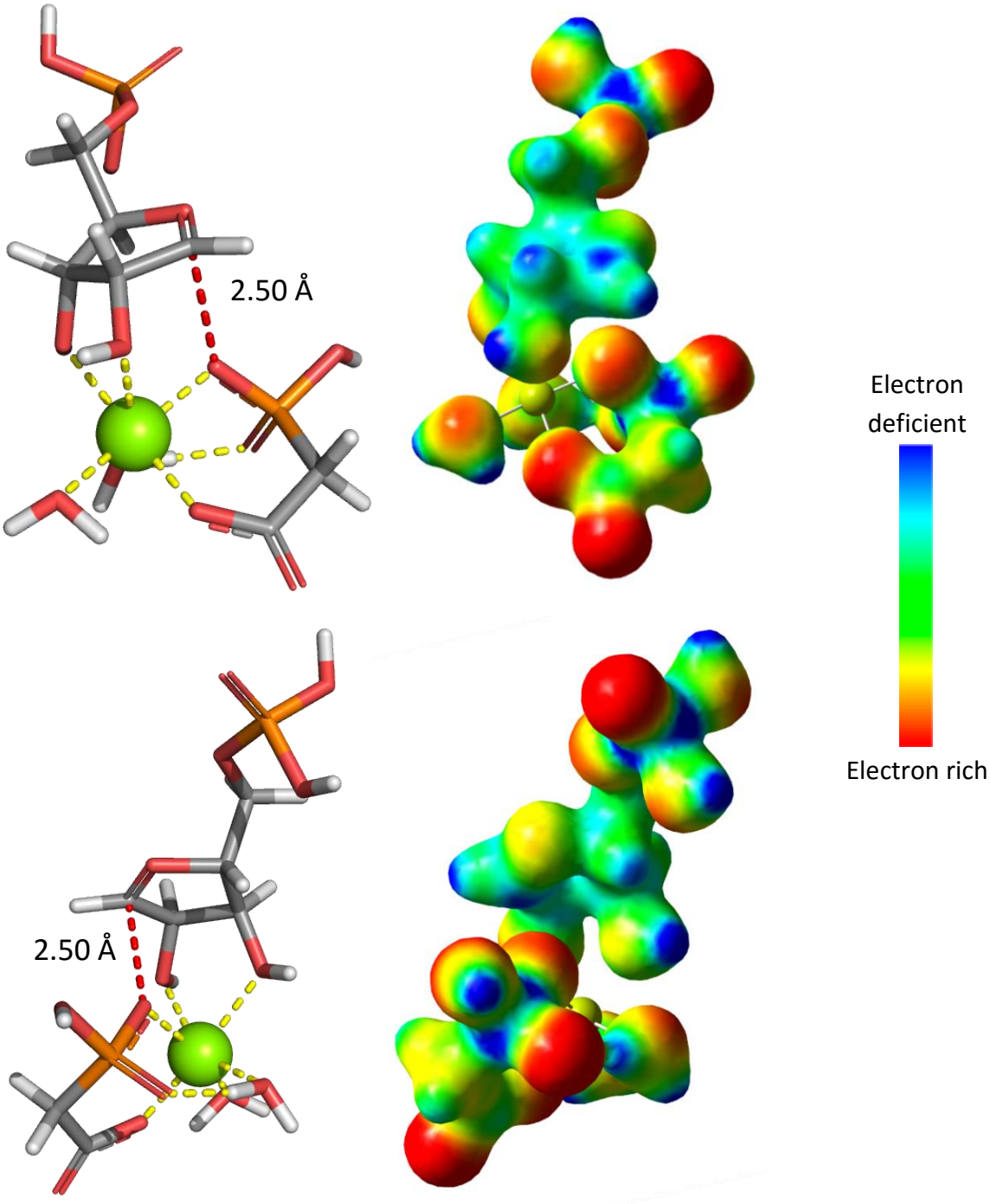


Figure 4.17. The transition state structure of L. lactis ATP-PRTase and the corresponding MEPS map. Scale on the side gives the arbitrary scaling with electron rich indicated in red and electron deficient in blue.

4.6 ATP-PRTase compared to other PRTase enzymes

Goitein *et al.* studied the primary ^{14}C and α-secondary ^{3}H KIEs for three phosphoribosyltransferase enzymes including ATP-PRTase from *Salmonella typhimurium*.[58] In their study commitment to catalysis suppressed the observed isotope effect in both forward and reverse directions (Section 3.2).[58] Human hypoxanthine guanine PRTase (HGPRTase) gave a primary isotope effect of 1.05 and an α-secondary isotope effect of 1.30. Yeast orotate PRTase (OPRTase) gave a primary KIE of 1.03 and an α-secondary isotope effect of 1.17. A more detailed transition state analysis based on experimental KIEs has also been carried out for OPRTase from *Salmonella typhimurium* by Tao *et al.* and *Plasmodium falciparum*, and human orotate phosphoribosyltransferase by Zhang *et al.* through the use of the alternative substrate phosphonoacetic acid.[60,61] OPRTase from *S. typhimurium* gave a primary KIE of 1.040 ± 0.004 and a α-secondary KIE 1.200 ± 0.007 supporting a $D_N{*}A_N$ transition state.[60] A relatively low nitrogen isotope effect was observed for double isotopically labelled orotate with a KIE of 1.006 ± 0.005. The transition state was modelled with a C1'-N1 bond distance of 1.85 Å and no involvement of the incoming nucleophile. This corresponds to a $D_N^{\ddagger}{*}A_N$ or a late dissociative reaction mechanism. *P. falciparum* and human OPRTase gave smaller primary isotope effects of 1.034 ± 0.002 for *P. falciparum* and 1.035 ± 0.003 for human OPRTase. The smaller primary isotope effect was supported by large α-secondary isotope effects of 1.261 ± 0.014 and 1.199 ± 0.015. In addition large N1 isotope effects of 1.028 ± 0.003 and 1.025 ± 0.005 were reported.[61] The transition state that agreed best with the experimental intrinsic KIEs had a fully dissociated dianionic orotate and a C1'-O^{PA} bond distance of 2.15 for *P. falciparum* and 2.14 for *H. sapiens* OPRTase which corresponds to a $D_N{*}A_N^{\ddagger}$ or late associative transition state. *P. falciparum* and *H. sapiens* orotate OPRTase has also been studied by Tao *et al.* with the natural substrate pyrophosphate. For this, commitment to catalysis was overcome through the use of orotidine instead of the natural substrate orotidine monophosphate (OMP) as a slow substrate. Also here a $D_N{*}A_N^{\ddagger}$ type mechanism was found, although a slight increase in the nucleophile bond distance to the anomeric position was modelled.

The transition state models for *P. falciparum* and *H. sapiens* OPRTase matches closest to the results found for ATP-PRTase. Both exhibit fully dissociated nitrogenous bases and a $D_N{*}A_N^{\ddagger}$ type mechanism based on the transition state analysis. The transition state model of *P.*

falciparum and *H. sapiens* OPRTase studied with PA however did not include magnesium at the transition state and was found to have a significantly shorter C1'-O^{PA} bond distance. In addition the PA nucleophile was modelled fully protonated. Based on the pK_a profile of PA it is likely that the carboxylic acid would be deprotonated similar to the transition state models proposed in this study. Deprotonation of the PA had a significant effect on the model used here and it may be these changes that led to significant differences between the models.

4.7 Long versus short; comparison of the two forms of ATP-PRTase enzyme

Previously significant differences between the transition state of human and bovine purine nucleoside phosphorylase (PNP) have been reported.[104,123,141] Bovine PNP goes through an early transition state with a distance of 1.8 Å between the C1' and N9. Human PNP on the other hand was determined to resemble a fully dissociated carbocation with a distances of ≥3.0 Å for the leaving group and nucleophile to the C1'. These are remarkable differences for enzymes with 87% sequence identity. Further study showed that a single point mutation of the Human PNP enzyme could change the transition state to resemble that of the bovine PNP.[142] Comparatively, *P. falciparum* and *H. sapiens* OPRTase share only 26% identity in amino acid sequence but display very similar transition states. *S. typhimurium* OPRTase on the other hand was shown to go through a $D_N^{\ddagger}*A_N$ or a late dissociative reaction mechanism.

These examples show that variation among the transition states for enzymes catalysing the same reaction is not uncommon. Although these enzymes are similar structurally, and in some cases in amino acid sequence, this does not necessarily lead to similar transition states. The ATP-PRTase long and short form are different from the given examples in that they share a common catalytic core but have vastly different means of regulation. Overall the ATP-PRTase enzymes studied share ~30% sequence identity, regardless of the difference in structure. The transition states for *M. tuberculosis*, *C. jejuni*, and *L. lactis* ATP-PRTase were found to be very similar with C1'-O^{PA} bond distances within 0.25 Å across the three enzymes and all three enzymes going through a $D_N*A_N^{\ddagger}$ or a late associative transition state. The close resemblance of the transition states for ATP-PRTase long form and ATP-PRTase short form support the findings that the regulatory domains are recruited for control of the complex

metabolic systems of cellular organisms and do not affect the reaction mechanism. These findings support the idea that both forms of ATP-PRTase share a common ancestor and that through divergent evolution the challenge of regulation was overcome for this metabolically important enzyme.

4.8 The reaction coordinate for ATP-PRTase

As mentioned in Section 4.3, experimental KIEs report on the difference between the substrate free in solution to the highest energy barrier along the reaction coordinate, the transition state. Intrinsic KIEs do not directly report on the steps leading to the formation of the TS. The TS models do however allow classification of a reaction (Section 4.2) and provide insight into the potential steps leading towards the transition state. In addition, Smith predicted a reaction mechanism through "nucleophilic displacement by electrophile migration".[112] This prediction was made from crystal structures with substrate, substrate analogues and transition state analogues bound at the active site of human HGPRT. It was also found that the nitrogenous base and pyrophosphate were held in place with most of the movement coming from the central atom, the anomeric carbon C1'.[143]

These findings support that the pyrophosphate and adenine are held in place at the active site of ATP-PRTases, with most of the movement during catalysis arising from migration of the central atom along the reaction coordinate. Depending on the interatomic distance between the nucleophilic oxygen on the PA phosphonate and the N1 of adenine, small variation in the transition state and the reaction coordinate may be found. This could help to further explain the subtle differences between the studied enzymes.

The three ATP-PRTase enzymes were modelled with similar transition states based on the primary ^{14}C and α-secondary ^{3}H. A bigger difference was found around the isotope effects associated with the adenine base 1-^{15}N and 6-^{15}N. The variation in active site may give rise to these observed differences and as a result the steps leading to the transition state could vary between the studied ATP-PRTase enzymes.

Large 1-^{15}N isotope effects for *M. tuberculosis* and *L. lactis* support the suggested $D_N{*}A_N^{\ddagger}$ reaction mechanism with ATP fully dissociated at the transition state or a bond distance ≥3.0 Å and weak interaction of the incoming nucleophile. A $D_N{*}A_N^{\ddagger}$ reaction mechanism suggests the presence of an unstable carbocation intermediate or TS. However, the KIE

results do not provide direct proof of this. Another explanation could be that the transition state is reached without going through this discrete intermediate and transition between dissociation of the nucleophile (1st TS) at a late stage is complemented by weak interaction of the incoming nucleophile (2nd TS), without formation of a discrete intermediate (Figure **A** 4.18).

C. jejuni ATP-PRTase on the other hand gave a primary nitrogen isotope effect near unity. An isotope effect at unity may be a result of the adenine dissociation step (1st TS) being further separated from the transition state on the reaction (2nd TS) (Figure **B** 4.18). The separation would allow for equilibration and restoration of the aromatic ring nullifying the KIE. This would further suggest that formation of a carbocation intermediate proceeds before the transition state of the reaction is reached (Figure **B** 4.18).

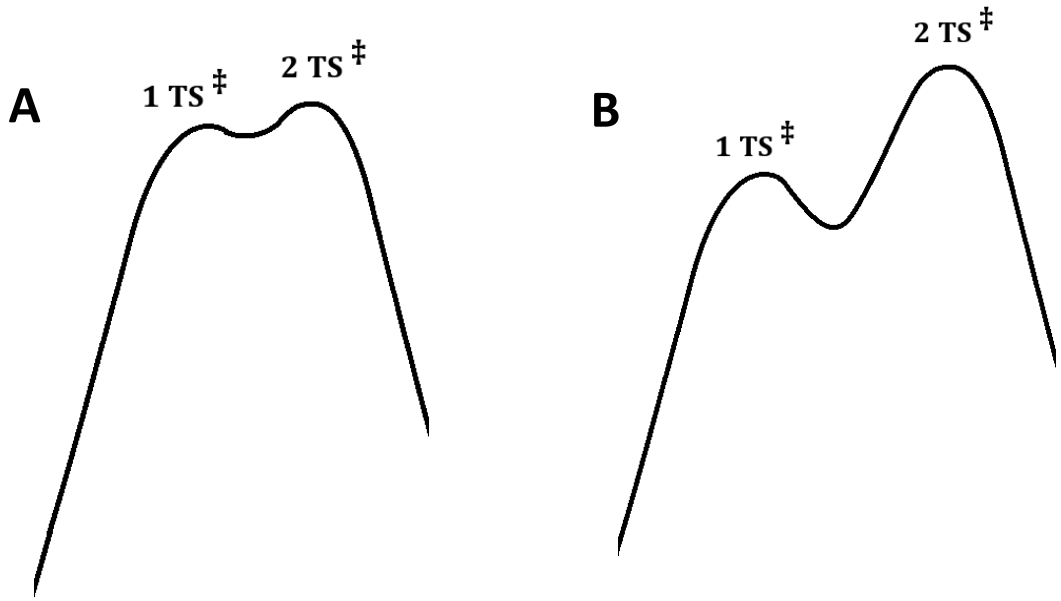


*Figure 4.18. Proposed difference in the reaction coordinate diagram between **A** the M. tuberculosis and L. lactis ATP-PRTase with dissociation of the nitrogenous base and incoming of the nucleophile without going through a discrete intermediate and **B** the reaction coordinate for C. jejuni ATP-PRTase with adenine dissociation (1st TS) separated by formation of an intermediate (carbocation) followed by incoming of the nucleophile (2nd TS).*

4.9 Summary

This chapter covered the generation of transition state models with the experimental intrinsic KIEs as boundary conditions. The ground state structure was successfully generated starting from PRATP bound to the *C. jejuni* ATP-PRTase *core* mutant. The full length structure was truncated at the ribose triphosphate which was shown to not adversely affect the calculated KIEs more than three bonds away. The nitrogen isotope effects for 1-^{15}N and 6-^{15}N were calculated for different charge and tautomeric states of the adenine ring system. The large 1-^{15}N isotope effect for ATP-PRTase from *M. tuberculosis* in combination with the residues at the active site make a neutral N3 and N7 tautomer possible, with the N3 tautomer considered the most likely because of a coordinated water in close proximity to this position. *L. lactis* ATP-PRTase was also found to have a fully dissociated adenine base, here in addition to the neutral N3 and N7 tautomers of adenine an anionic leaving group could not be ruled out with additional interactions to adenine proposed at the active site. The N7 tautomer was considered a more likely candidate, as an anionic adenine ring would form a poor leaving group. In addition, in the physiological reaction order this would require deprotonation of the adenine amine. Given the pK_a of ~16 at this position, this was considered less likely to occur even if multiple interactions would facilitate this step. *C. jejuni* ATP-PRTase was considered to have a fully dissociated adenine based on the findings of the other two enzymes, and that the transition state resembles the PRPA product. The high level of conservation for the PRPP binding loop across phosphoribosyl transferase enzymes allowed PRPP to be used as a starting structure and the generation of the PRPA transition state model. The ribose magnesium complex was refined separately and the C1'-O^{PA} bond incrementally elongated. The 5'-phosphate was then built back into the model and refined further to give the transition states models. To study the effect of different functionals a grid search was carried out with the B3LYP functional performing best out of the functionals tested. A C1'-O^{PA} bond length of 2.75 Å gave the best match to the experimental data of *M. tuberculosis* ATP-PRTase. *C. jejuni* ATP-PRTase gave a very similar result with a bond length of 2.65 Å. Also the ATP-PRTase short form had a transition state with a C1'-O^{PA} bond length of 2.50 Å. This indicates that all three enzymes go through a late associative $D_N{*}A_N^{\ddagger}$ transition state. Magnesium plays an important role in anchoring the reactants in place during transition state formation by coordinating the reactants in an octahedral configuration. The carbocation is stabilised by the ribose ring oxygen which

shares the electrons of a filled lone-pair orbital with an empty *p*-orbital of the carbocation. This results in a double bond characteristic which has a ring flattening effect on the ribose ring. The work presented revealed a similar transition state for the three enzymes and supports the proposed reaction mechanism that the carbocation drives the reaction with "nucleophilic displacement by electrophile migration". Small differences are likely to be found in the steps leading to the transition state particularly around the adenine base and leaving group departure and that the nature of the transition states is determined by the catalytic site architecture.

Chapter 5

Synthesis of transition state analogues

5.1 Tight binding of transition state analogues

The original proposal by Pauling that enzymes bind tightly to the transition state was based on transition state theory (TST) which states that reactions are in equilibrium with the transition state.[144] Pauling proposed that the catalytic potential of an enzyme was achieved through tight binding, and therefore stabilisation, of the transition state complex to the enzyme. Based on this and the notion that reaction rate is proportional to the fraction of the reactant in the transition state complex, Wolfenden and co-workers further hypothesized that enzymes are able to alter this equilibrium through specific tight binding to the transition state therefore providing stabilisation of the transition state complex.[145,146] The binding affinity of the enzyme for the transition state was proposed to be proportionate to the rate increase, of 10^{10} to 10^{20} times the uncatalysed rate, imposed by the enzyme.[147] An inhibitor that can capture some of the nature of the transition state geometry, van der Waals interactions and charge are predicted to bind tightly to the enzyme, with the binding energy of the transition state analogue proportional to the catalytic rate enhancement.[145]

Examples of tight binding, when enough of the transition state properties are captured, are well reported.[141,148–150] Some of the transition state mimics bind with staggering picomolar to femtomolar affinity to their target. Furthermore, in the case of purine nucleoside phosphorylase (PNP) it was shown that bovine and human PNP have different transition states. The difference in transition state could be exploited in transition state analogue design by insertion of a methylene bridge between the ribose- and purine- mimicking moyity. The enzyme exhibiting a more dissociated transition state was able to bind the extended transition state analogue more tightly then the enzyme displaying an earlier transition state.[2,123,141]

This chapter discusses work towards finding alternative means to target the ATP-PRTase enzymes of pathogenic bacteria, in particular the ATP-PRTase enzymes from *M. tuberculosis* and *C. jejuni*. Work involving the design and synthesis of transition state analogues as a means to inhibit these enzymes is described.

5.2 Transition state analogue design

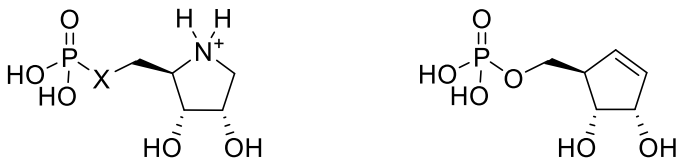
Enzymes with an altered geometry and/or charge at the transition state are good candidates for transition state analogue design.[108] At the start of this work a detailed description of the

transition state structure was unavailable. Based on crystal structures and work done on other phosphoribosyl transferases the ATP-PRTase enzyme was predicted to follow an S_N1 like transition state (Section 1.11). This suggests a carbocationic species with charge build up around the anomeric centre and an altered geometry at the site of the reaction, from sp^3 or tetrahedral geometry, to more sp^2 trigonal planar geometry.

Knowledge of the nature of the transition state allows for more informed design of inhibitors, leading to the very potent inhibitors mentioned earlier. Synthesis of these inhibitors, however still remains challenging and time consuming. To this end inhibitors based around the ribocation portion were chosen first, before embarking on a full length PRATP based analogue.

Iminosugars are sugar derivatives with the ring oxygen replaced by nitrogen. These are protonated at physiological pH and represent the positive charge of the transition state. It is these particular species that have been shown to bind with such high potency, and based on this, targets **5.1** and **5.2** were chosen (Figure 5.1). The phosphate group plays an important role in binding PRPP, so it was therefore included. In addition, the phosphonate isostere was chosen for the iminosugar derivative as a stable analogue of the phosphate group which is easily accessible through similar intermediates as compound **5.1**.

The positive charge, built up at the anomeric carbon during ribocation formation, is stabilised by the ring oxygen through its lone pair electrons. This results in double bond characteristics and the trigonal planar species. To capture the geometry of the flattened ribose ring cyclopentene **5.3** was chosen to be synthesised (Figure 5.1).



5.1: X = O
5.2: X = CH_2

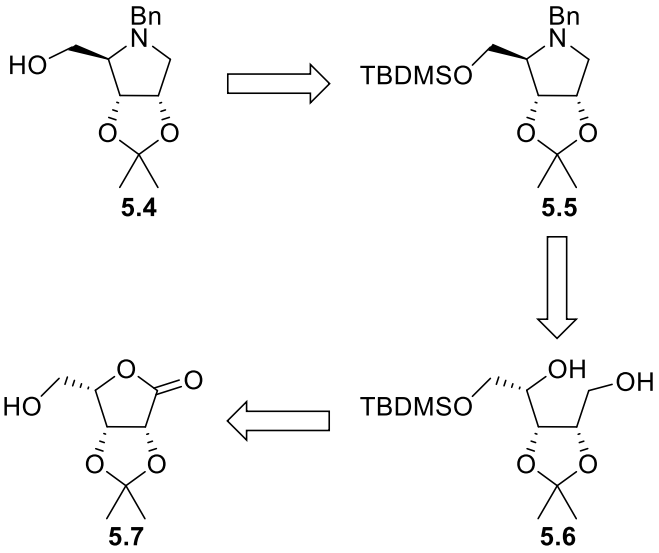
5.3

*Figure 5.1. The proposed intermediates: iminosugars **5.1** and **5.2** as a mimic of the ribocation intermediate and the cyclopentene **5.3** as a mimic of the ring flattening effect.*

5.3 Iminoribitol transition state analogue

5.3.1 Retrosynthetic analysis of the shared iminosugar portion.

Both phosphate **5.1** and phosphonate **5.2** were envisioned to be accessible from the shared intermediate hydroxymethyl **5.4** which, in turn, was envisioned to be accessed in five synthetic steps (Scheme 5.1). It was proposed that orthogonally protected hydroxymethyl containing iminosugar **5.4** could be obtained in one step from fully protected iminoribitol **5.5** by desilylation. Silane **5.5** in turn was expected to be accessible in two steps from diol **5.6** by installation of the mesylate leaving group followed by a double displacement with benzylamine and inversion of the stereochemistry at the C4-position. Diol **5.6** itself was thought to be accessible from lactone **5.7** via silane protection of the 5-position followed by reduction of the lactone portion.

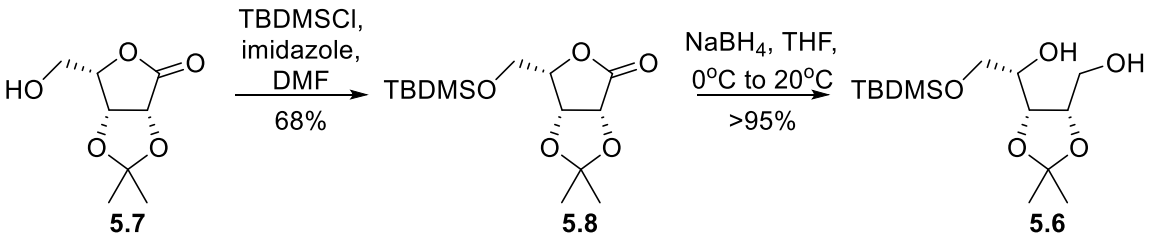


Scheme 5.1. Retrosynthetic analysis of iminosugar **5**.

5.3.2 Synthesis of the iminosugar portion

The starting lactone was kindly provided by Prof. P. Tyler from the Ferrier Research Institute, Wellington NZ. The synthesis commenced with protection of the 5-position of the starting material by subjecting lactone **5.7** to *tert*-butyldimethylsilyl chloride (TBDMSCl) and imidazole in dimethylformamide (DMF) (Scheme 5.2). This resulted in the unexpected appearance of two products by thin layer chromatography (TLC), with one of these later confirmed to come from an impurity in the starting material carried over from the previous

step. The product could readily be separated from the unwanted side product by flash chromatography and identity of the desired product was confirmed by X-ray crystallography with a flack value of 0.01 ± 0.03 (Figure 5.2, appendix IV). Subsequent reduction proceeded smoothly by subjecting protected lactone **5.8** to sodium borohydride in tetrahydrofuran (THF). NMR analysis suggested full conversion and no apparent side product; therefore, diol **5.6** was used without further purification.



*Scheme 5.2. Synthesis of diol **5.6** from lactone **5.7** in two steps via protection of the primary hydroxyl followed by reduction with $NaBH_4$.*

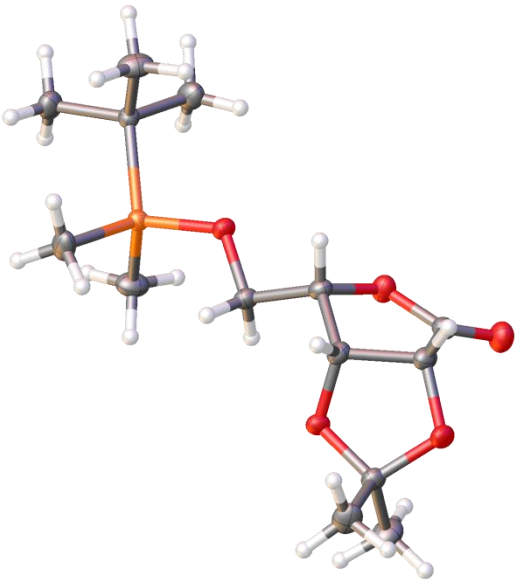
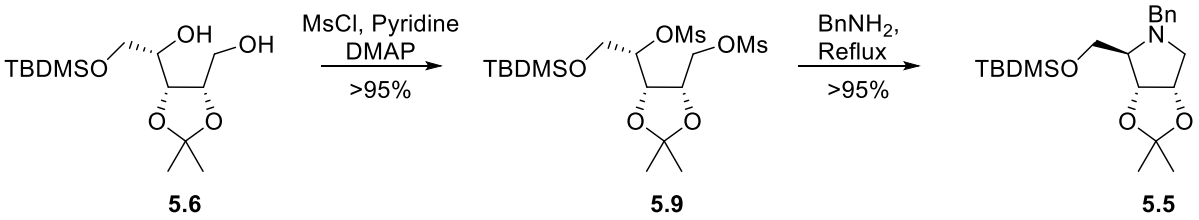


Figure 5.2. The crystal structure of lactone 5.8 confirming the right stereochemistry and product.

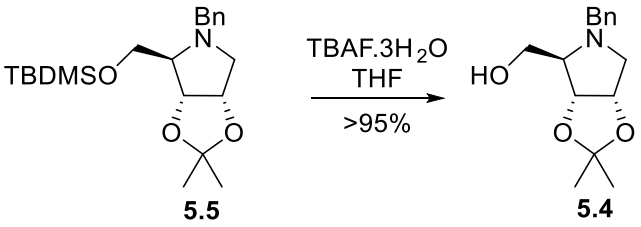
Following a modified procedure of Fleet *et al.*, diol **5.6** was subjected to methanesulphonyl chloride (MsCl) and DMAP in pyridine (Scheme 5.3).[151] In this reaction it was found that a better yield was obtained when MsCl was activated with DMAP prior to the addition of diol **5.6**. Following the adopted method, good formation of dimesyl **5.9** was observed (>95%). To perform the double displacement, dimesyl **5.9** was dissolved in benzyl amine and heated to

160 °C for four hours. This gave clean conversion of the starting material to the protected amine product **5.5,** in excellent yield (>95%).



*Scheme 5.3. Installation of the mesyl leaving group followed by a double displacement with benzyl amine to give iminosugar **5.5**.*

The next step in the synthesis was the deprotection at the 5-position of orthogonally protected amine **5.5** for subsequent modifications around this centre. Therefore amine **5.5** was subjected to tetrabutylammonium fluoride trihydrate in THF, to give primary alcohol **5.4** in near quantitative yield (Scheme 5.4).



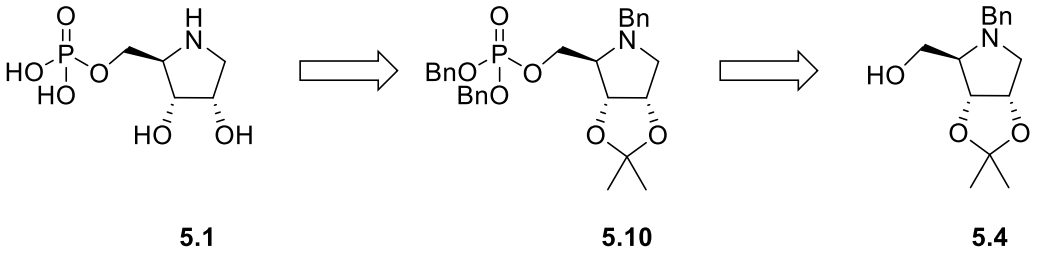
*Scheme 5.4. Deprotection of **5.5** to give orthogonally protected primary alcohol **5.4**.*

5.3.3 Synthesis of the iminoribitol phosphate

5.3.3.1 Retrosynthetic analysis of iminoribitol phosphate

The desired iminoribitol phosphate **5.1** was expected to be accessible in three steps from primary alcohol **5.4** (Scheme 5.5). Global deprotection of fully protected iminoribitol **5.10** was expected to be possible in two steps through acid catalysed hydrolysis of the acetonide followed by hydrogenation. Installation of the phosphate functional group itself was envisioned to be accessible in one step from alcohol **5.4** in a two-step one-pot procedure by phosphorus coupling with phophoramidite and a weak acid followed by oxidation of the

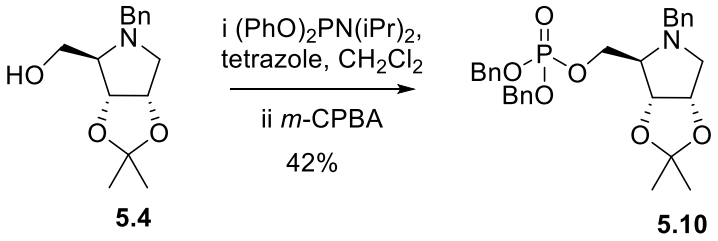
phosphite group. The benzyl protecting group was chosen to allow for a convergent deprotection strategy.



*Scheme 5.5. Retrosynthetic analysis of iminoribitol phosphate **5.1**.*

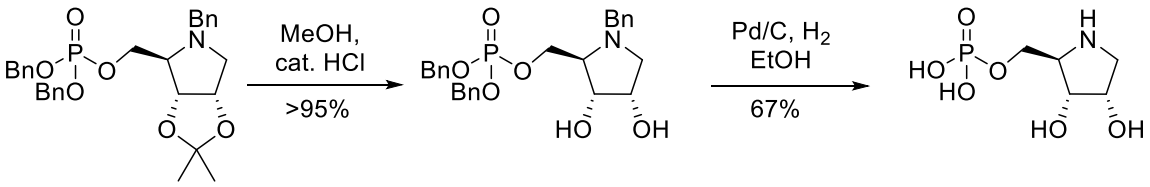
5.3.3.2 Synthesis of the iminoribitol phosphate

The hydroxymethyl **5.4** was subjected to dibenzyl *N*,*N*-di-isopropyl-phosphoramidite and tetrazole (Scheme 5.6).[152] When TLC analysis showed complete consumption of starting material *meta*-chloroperoxybenzoic acid (*m*-CPBA) was added to give protected phosphate **5.10** in a moderate 42% yield.



*Scheme 5.6. Phosphorylation of hydroxymethyl **5.4**.*

Although formed in modest yield, the amount of protected iminoribitol phosphate **5.10** was considered sufficient for the two step deprotection (Scheme 5.7). Following a method by Zhang *et al.* hydrolysis of the acetonite protecting group was readily achieved in methanol and a catalytic amount of HCl to give the 2,3-dihydroxy product **5.11** in a good yield (>95%).[150] In this step longer reaction times resulted in the loss of one of the *O*-benzyl protecting groups as well. Subsequent hydrogenation then gave the desired product iminoribitol **5.1** in a total of 8 steps and an overall yield of 15% starting from lactone **5.7**.



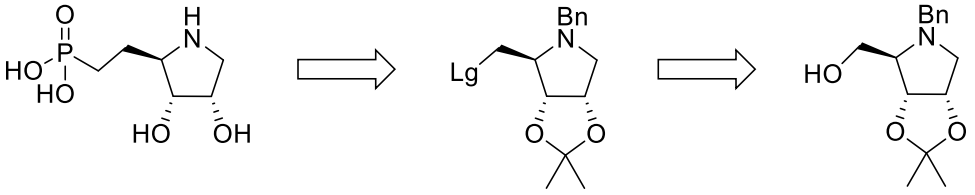
5.10 **5.11** **5.1**

*Scheme 5.7. Deprotection of iminoribitol **5.10** via hydrolysis of the acetonite protecting group followed by hydrogenation.*

5.3.4 Synthesis of the iminoribitol phosphonate transition state analogue

5.3.4.1 Retrosynthetic analysis of iminoribitol phosphonate

Phosphonate **5.2** was envisioned to be accessible in four synthetic steps from the shared hydroxymethyl **5.4** (Scheme 5.8). Introduction of the phosphonate was anticipated to be possible by introducing a good leaving group (**5.13**) at the 5-position and subsequently displaced by a phosphonate based nucleophile. Deprotection was expected to be possible in a two-step process of acid hydrolysis followed by hydrogenation.



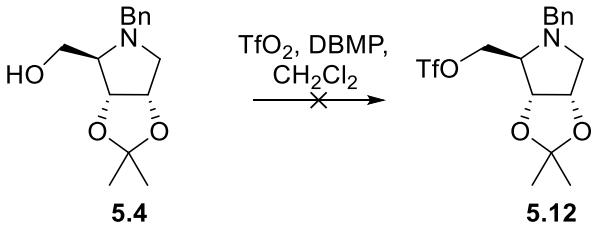
5.2 **5.13** **5.4**

*Scheme 5.8. retrosynthetic analysis of iminoribitol phosphonate **5.2**.*

5.3.4.2 Synthesis of iminoribitol phosphonate

The trifluoromethansulphonyl (Tf) leaving group was chosen as the leaving group because of the highly activated nature and high reactivity towards nucleophilic displacement. Primary alcohol **5.4** was therefore subjected to trifluoromethanesulphonic anhydride and 2,6-di-*tert*-butyl-4-methylpyridine (DBMP) as a base (Scheme 5.9). However, analysis by NMR and mass spectrometry revealed that only eliminated product was formed. Formation of the eliminated product from triflate **5.12** was attributed to the activated nature of the triflate leaving group. A less activated leaving group, such as an iodide or methyl sulphonyl, could potentially reduce the formation of the eliminated product and their synthesis was

attempted. Installing the mesylate resulted in the formation of multiple side products during the reaction and was not followed up further. Mass spectrometry analysis on the iodide gave a product with the correct mass, as well as a small amount of an eliminated product. Structural analysis by NMR spectroscopy, however did not match the expected product for the iodide with a peak at 29 ppm corresponding to a secondary iodide by the HSQC spectrum (Figure 5.3).



5.4 **5.12**

Scheme 5.9. The installation of a triflate leaving group did not yield the desired product

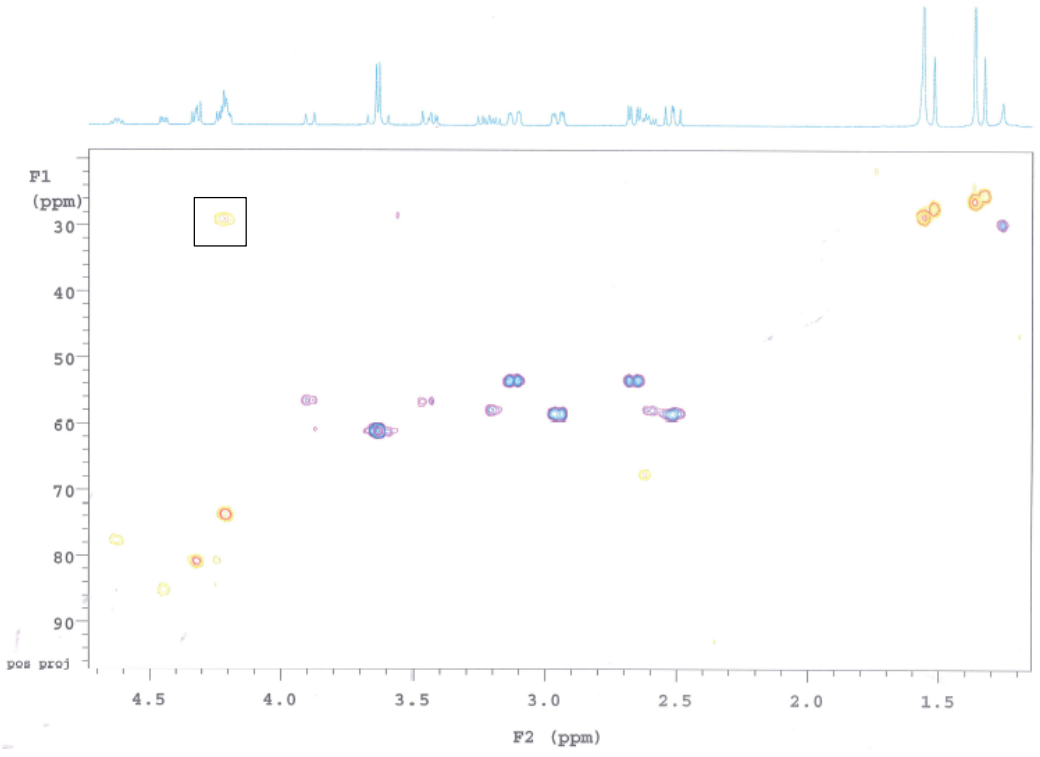
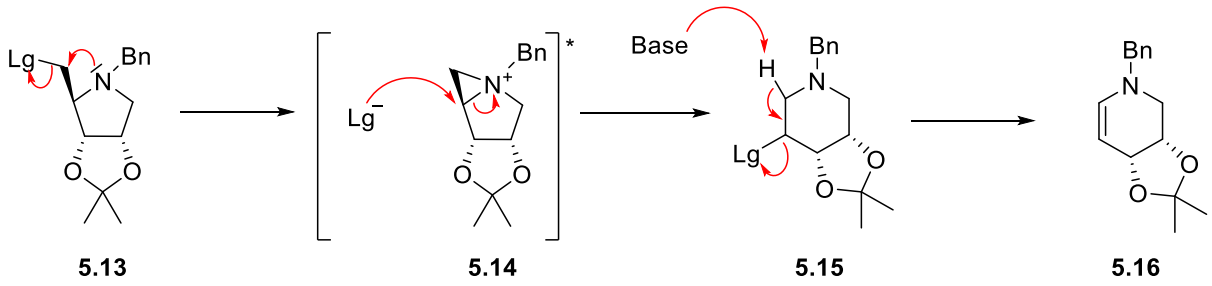


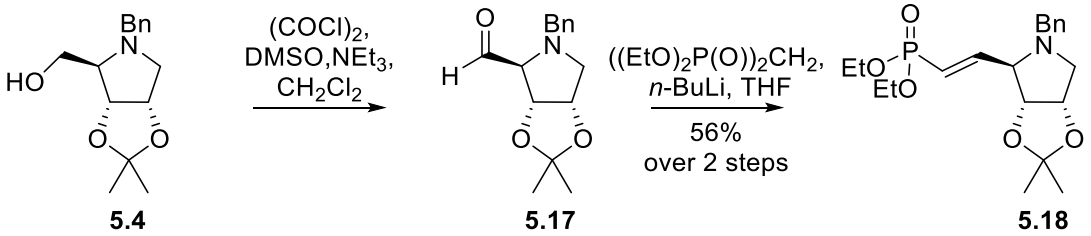
Figure 5.3. [^{1}H,^{13}C]-HSQC spectrum of the secondary iodide with all CH and CH_3 centres in yellow and the CH_2 centres in blue. The product gave a CH centre at 29 ppm supporting a secondary iodide.

Formation of a secondary iodide can be explained through the participation of the ring nitrogen which sits in close proximity to the leaving group (Scheme 5.10). Mechanistically this can be explained by the ability of the filled orbital of the lone pair on nitrogen to displace the installed leaving group, resulting in a positively charged aziridine intermediate (**5.14**). Nucleophilic attack on this strained ring system by the released leaving group then results in ring expansion to give the six membered ring **5.15**. The participation of the ring nitrogen is further improved by the electron donating effect of the benzyl protecting group. Ring expansion was also observed by Malladi *et al.* when attempting to introduce the mesyl leaving group.[153] β-elimination of **5.15** then leads to alkene **5.16**.



*Scheme 5.10. Formation of the ring expanded product **5.15** and subsequent β-elimination leading to alkene **5.16**.*

To circumvent the problem around the ring nitrogen participating in the reaction, two approaches were considered. If the centre of interest would be oxidised instead, participation by the ring nitrogen was not expected to lead to the discussed side products. Another option would be to change the nitrogen protecting group to *tert*-butyloxycarbonyl (Boc). The formed carbanate has an electron withdrawing effect on the nitrogen, thus reducing the nucleophilicity of the nitrogen. The oxidation method was the preferred option as this would not introduce additional steps. Alcohol **5.4** was therefore subjected to a Swern oxidation to give aldehyde **5.17** and dimethylsulphide (Me_2S) as the only observed products by NMR in the crude reaction (Scheme 5.11). The aldehyde was used without further purification to prevent side product formation observed during purification by silica flash chromatography. Treatment of aldehyde **5.17** under HWE conditions led to product formation and gave phosphonate **5.18** in a modest 56% yield over two steps.



*Scheme 5.11. Installation of the phosphonate functionality via the aldehyde **5.17** to give phosphonate **5.18**.*

With the fully protected phosphonate **5.18** in hand, deprotection was attempted (Scheme 5.12). Hydrolysis of the ethyl esters was the first step in the deprotection regimen, to this end phosphonate **5.18** was subjected to TMSBr in CH_2Cl_2. This resulted in clean formation of the free acid. Subsequent removal of the isopropylidine was then achieved by acid catalysed hydrolysis with acetic acid/water 1/1 v/v at 70 °C. When clean conversion was confirmed by mass spectrometry the crude reaction mixture was subjected to 10% palladium on carbon with hydrogen gas in methanol to give iminoribitol phosphonate **5.2** in 83% yield over three steps after purification. Overall this gave the phosphonate **5.2** in 8 steps with an overall yield of 26%.

*Scheme 5.12. Deprotection of iminoribitol phosphonate **5.18**.*

5.3.5 Inhibition studies on iminoribitol based analogues.

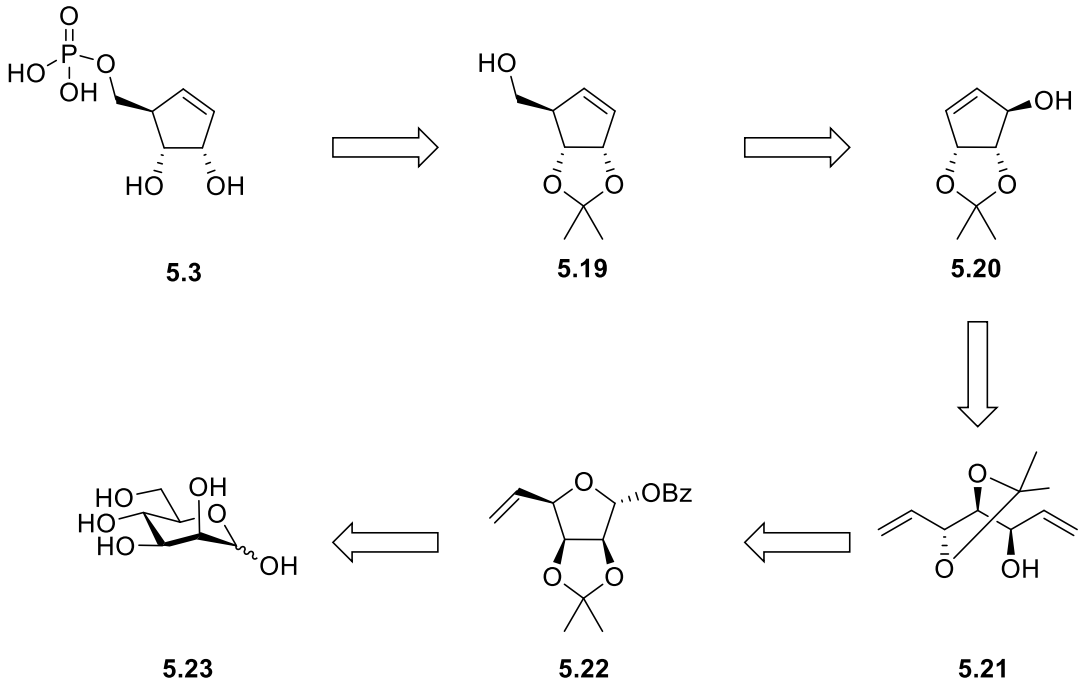
The iminoribitol phosphate **5.1** and iminoribitol phosphonate **5.2** were tested under competitive conditions (Section 8.5.3). The inhibitors were expected to be competitive towards PRPP binding; therefore, the concentration of PRPP was varied (1–3 times K_m: 50-150 µM for the *M. tuberculosis* ATP-PRTase enzyme[70] and 11–33 µM for the *C. jejuni* ATP-PRTase enzyme[66]) while keeping ATP concentrations constant (750 µM *M. tuberculosis* ATP-PRTase and 500 µM *C. jejuni* ATP-PRTase enzyme), at an inhibitor concentration of up to 500 µM.

No inhibition for the ATP-PRTase enzyme from *M. tuberculosis* and *C. jejuni* under these conditions was observed for either analogue. The iminoribitol based inhibitors mimic the ribose-5-phosphate moiety which is only a small part of the full length PRATP structure able to bind at the active site. Potentially, the adenine base or a similar moiety may be needed to allow binding of the TS analogues in the right geometry to effectively inhibit these enzymes.[141]

5.4 Ribose cyclopentene analogue

5.4.1 Retrosynthetic analysis

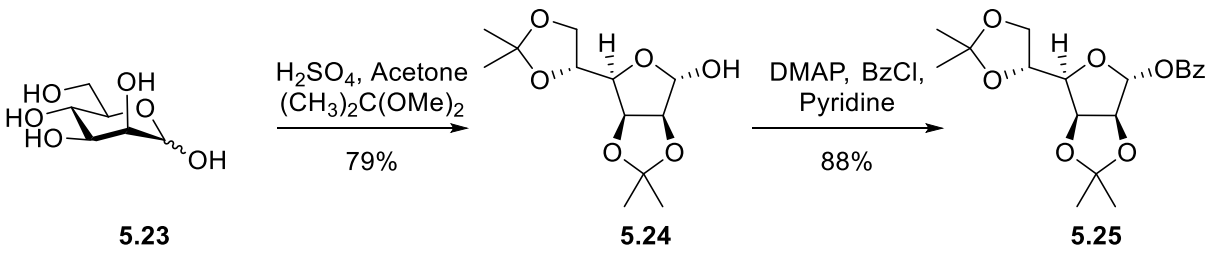
A scheme for the synthesis of cyclopentene **5.3** was devised and comprises ten steps starting from D-mannose **5.23** (Scheme 5.13). It was envisioned that phosphorylated cyclopentene **5.3** could be obtained from protected cyclopentene **5.19** in two steps via a protected phosphoryl chloride and an organic base followed by global deprotection. Following a literature procedure by Overkleeft *et al.* cyclopentene **5.19** itself has been synthesised in eight steps.[154] Cyclopentene **5.19** was synthesised through the condensation of tributyl(iodomethyl)stannane with cyclopentene **5.20** followed by a Still-Wittig rearrangement. Cyclopentene **5.20** was synthesised from dialkene **5.21** using ring closing metathesis (RCM) with Grubbs 1st generation catalyst. Dialkene **5.21** in turn was obtained from orthogonally protected alkene **5.22** via de-benzoylation and consecutive Wittig olefination. Alkene **5.22** was obtained via protecting group manipulations and acid catalysed thermal rearrangement from readily available D-mannose **5.23**.



*Scheme 5.13. Retrosynthetic analysis for the synthesis of the cyclopentene **5.3**.*

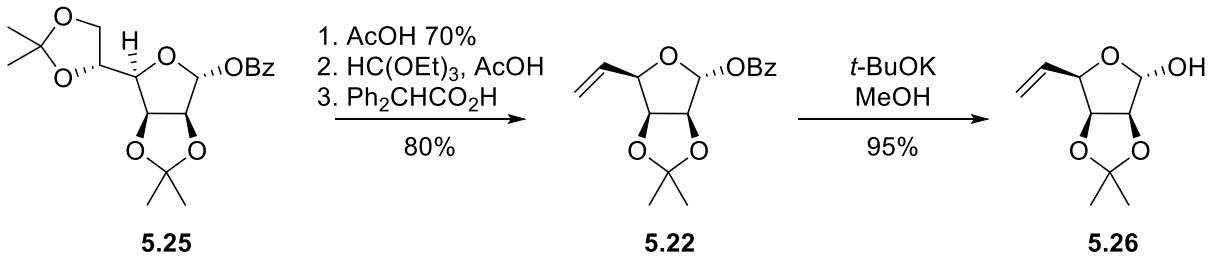
5.4.2 Synthesis of the cyclopentene transition state analogue

Protected mannose **5.25** was obtained in 2-steps from D-mannose **5.23** (Scheme 5.14). The acetonide protecting group was chosen to favour the mannose ring in its furanose form and the ability to modify the 5,6 isopropylidine protecting group selectively at a later stage in the synthesis. D-Mannose **5.23** was subjected to the condensation reaction with 2,2-dimethoxypropane in dry acetone and catalytic H_2SO_4 to give di-isopropylidine **5.24** in 79% yield. Benzoylation of the anomeric hydroxyl was then achieved with benzoyl chloride in dry pyridine and a catalytic amount of 4-dimethylaminopyridine (DMAP) to give fully protected mannose **5.25** in a good yield (88%).[155]



*Scheme 5.14. Synthesis of benzoyl 2,3:5,6-di-O-isopropylidine-α-D-mannose **12**.*

With the protected mannose **5.25** in hand, attempts were made at the selective deprotection, reprotection and thermal rearrangement at the 5,6 position to give protected alkene **5.22** (Scheme 5.15). Regio-selective removal of the 5,6-isopropylidine was achieved by treating protected mannose **5.25** with 70% acetic acid in H_2O at 45°C for five hours. It was important to monitor the progress of the reaction, as prolonged reaction time can lead to loss of the 2,3-isopropylidine as well. The crude product was then subjected to triethylorthoformate, followed by acid catalysed thermal rearrangement. In the literature Ph_3CCO_2H has been used for reaction at elevated temperatures as it is a less volatile acid catalyst.[156] Ph_3CCO_2H was unavailable at this point and therefore Ph_2CHCO_2H was used instead as this has similar properties and was present in the laboratory at the time. This gave protected alkene **5.22** in 80% yield over 3-steps. Subsequent deprotection of the benzoyl protecting group was readily achieved with *t*-BuOK in methanol to give lactol **5.26** in excellent yield (95%).

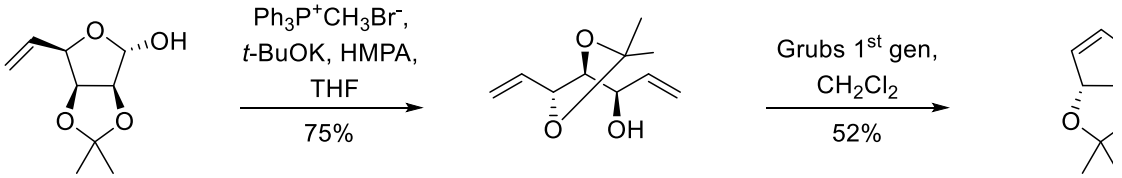


5.25 **5.22** **5.26**

*Scheme 5.15. Acid catalysed thermal rearrangement leading to alkene **5.22** followed by deprotection of the benzoyl protecting group to liberate the anomeric proton.*

With lactol **5.26** successfully synthesised, Wittig olefination was then explored. In an initial attempt, the lactol **5.26** was subjected to methyltriphenylphosphonium bromide and *n*-butyl lithium (*n*-BuLi) in THF .[157] Although product formation was observed over the course of the reaction (R_f 0.69, PE/EtOAc, 1/1, v/v), after purification no significant amount of product was isolated (<20%). A change of condition to *t*-BuOK in THF/ hexamethylphosphoramide was attempted in order to install the additional alkene (Scheme 5.16). However, this did not significantly improve the yield, although complete consumption of starting material to a higher running product was observed by TLC analysis. It was later realised that the dialkene product was volatile under reduced pressure and care had to be taken when drying the product under reduced pressure. When the reaction was repeated, bearing this in mind (vacuum >150 mbar and the water bath at 35°C), dialkene **5.21** was isolated in a yield of

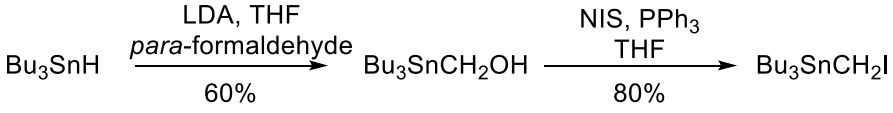
75%. RCM with Grubbs 1st generation catalyst then gave cycloalkene **5.20** in a 52% yield.[158] Although this yield was modest, enough of the cyclic product **5.20** was obtained to continue with the synthesis and was therefore not further optimised.



*Scheme 5.16. Wittig olefination followed by RCM with Grubs 1st generation catalyst leading to cyclopentene **5.20**.*

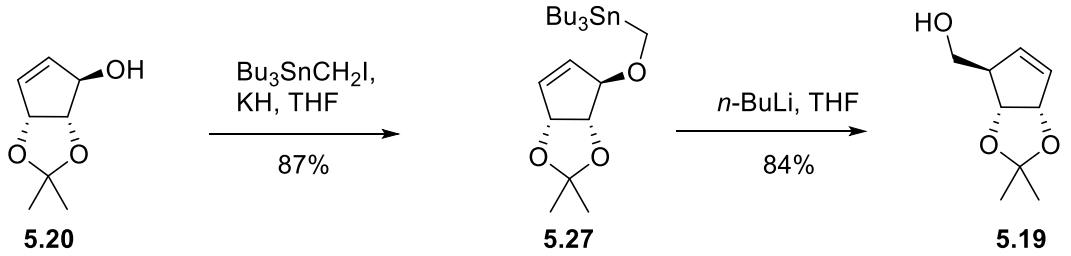
The next key step in the synthesis was the Still-Wittig [2,3] sigmatropic rearrangement. The Still-Wittig rearrangement is a variation of the traditional [2,3] Wittig sigmatropic rearrangement where the allylic ether is replaced by α-alkoxystannane, which after transmetalation with lithium can undergo the rearrangement step.

The tributyl(iodomethyl)stannane needed for this reaction was not readily commercially available and it was therefore synthesised. Following a procedure by Lohse-Fraefel *et al.* the tributyl(iodomethyl)stannane was synthesised in two steps (Scheme 5.17).[159] First tributyltin hydride was subjected to LDA and *para*-formaldehyde in THF leading to the formation of tributyl(hydroxylmethyl) stannane in 60% yield. The hydroxyl intermediate could be readily purified by flash chromatography. Subsequent iodonation with *N*-iodosuccinimide (NIS) and triphenylphosphine in THF then gave tributyl(iodomethyl)stannane in an 80% yield.



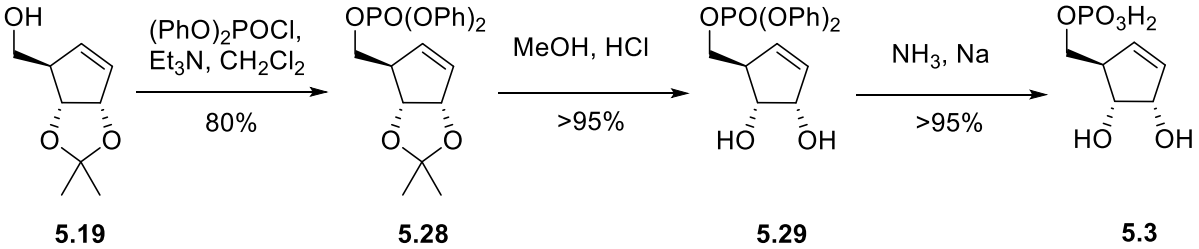
Scheme 5.17. Synthesis of tributyl(iodomethyl)stannane.

Starting from the iodomethyl stannane reagent, precursor **5.27** was then synthesised. Cyclopentene **5.20** was subjected to potassium hydride in THF, followed by the addition of tributyl(iodomethyl)stannane, to give ether **5.27** in an 87% yield (Scheme 5.18). Still-Wittig rearrangement was then achieved by subjecting precursor **5.27** to *n*-BuLi in THF to give the ribocarboncycle **5.19** in an 84% yield. In addition some of the hydrolysed product (**5.20**) was isolated.



Scheme 5.18. [2,3] Still-Wittig sigmatropic rearrangement.

The Still-Wittig rearrangement leaves the hydroxymethyl conveniently available for subsequent instalment of a phosphate group. The phenyl protected phosphate was chosen because it can be readily deprotected using a Birch reduction and avoids the use of TMSBr, which can be hard to remove in the final stages. Subjecting hydroxymethyl **5.19** to diphenylphosphoryl chloride and triethyl amine in dichloromethane gave the desired phosphate **5.28** in an 80% yield (Scheme 5.19). Removal of the acetonitrile protecting group was then achieved by subjecting **5.28** to concentrated HCl in methanol for two minutes followed by evaporation to dryness on a rotary evaporator. This gave diol **5.29** as the only observed product by NMR and was used without further purification. Subsequent deprotection by Birch reduction gave the desired phophocyclopentene **5.3** as the only observed product at >95% yield.



*Scheme 5.19. Phosphorylation and deprotection of phosphocyclopentene **5.3**.*

5.4.3 Inhibition study on cyclopentene phosphate

The IC50, or the concentration of cyclopentene phosphate **5.3** that provokes a response of half of the maximum inhibition, was measured for the *M. tuberculosis* ATP-PRTase enzyme. During these measurements the ability to inhibit catalysis was assessed at various concentration of the cyclophopentene phosphate inhibitor while the PRPP and ATP concentration were kept constant. It was found that the analogue was able to inhibit the *M.*

tuberculosis ATP-PRTase enzyme with an IC50 of 48 ± 4 µM (Figure 5.4). The IC50 was calculated according to Equation 14 employed in the graphpad prism 6.0 program.

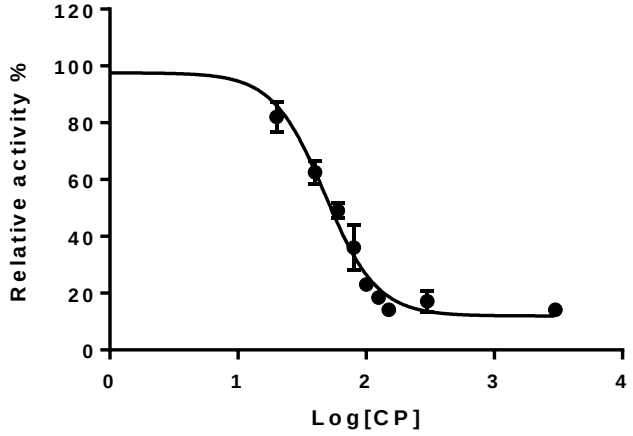


Figure 5.4. Dose response curve for the inhibitor 5.3 against M. tuberculosis ATP-PRTase. Error bars represent the error as the standard deviation with each datapoint from at least two independend measurements.

$$Y = Bottom + (Top - Bottom)/(1 + 10^{((logIC50 - X) * HillSlope)}) \quad \text{Equation 14}$$

Cyclopentene phosphate **5.3** was also tested for inhibition again the *C. jejuni* ATP-PRTase enzyme (Section 8.5.3). The inhibitors was expected to be competitive towards PRPP binding, therefore the concentration of PRPP was varied (1 to 3 times K_m: 11–33 µM)[67] while keeping ATP concentrations constant (500 µM)(Equation 15 and 16 and Figure 5.5). Cyclopentene phosphate **5.3** was found to be a moderate inhibitor for the C. jejuni ATP-PRTase enzyme with an K_i of 282 ± 66 µM.

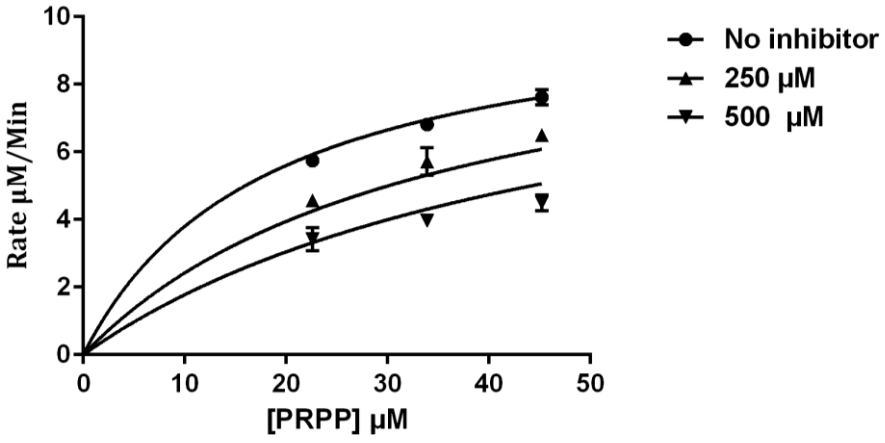


*Figure 5.5. Inhibition of C. jejuni ATP-PRTase by cyclopentene **5.3**. Inhibition was measured at varying PRPP concentration while keeping ATP constant. Error bars represent the error as the standard deviation with each datapoint from at least two independend measurements.*

$$K_{m_{observed}} = K_m * (1 + [I]/K_i) \quad \text{Equation 15}$$

$$Y = V_{max} * X/(K_{m_{observed}} + X) \quad \text{Equation 16}$$

5.5 Summary

The cyclopentene phosphate was successfully synthesised in eleven steps starting from D-mannose. This molecule was found to be a inhibitory for both enzymes with an IC50 of 48 ± 4 µM for the *M. tuberculosis* ATP-PRTase enzyme and an apparent K_i observed of 282 ± 66 µM for *C. jejuni* ATP-PRTase. These results illustrate that inhibition can be observed even if small components of the transition state are captured in an analogue. Several attempts at co-crystallising (inhibitor concentrations of up to 10mM) or soaking (up to 50 mM) did not result in the observation of clear electron density, but based on the structural similarity of the inhibitor to PRPP, the cyclopentene phosphate is expected to bind in the PRPP binding site. The phosphate group is expected to be accommodated by the phosphate binding loop and the cyclopentene ring hydroxyl groups to interacting with Glu141 and Asp154. The observation that the cyclopentene phosphate is a competitive inhibitor for PRPP suggests that it would form a good starting point for further inhibitor design. The incorporation of structural features of the second substrate ATP could be a good starting point for further development of this inhibitor (discussed in more detail in Section 7.3.1).

The iminoribitol phosphate and phosphonate were synthesised from a common intermediate in eight and ten steps respectively. It was found that both analogues were unable to inhibit the ATP-PRTase enzyme from *M. tuberculosis* and *C. jejuni* under the conditions tested. This suggests that this initial analogue was unable to mimic enough of either the substrates or transition state for effective binding. Given the ability of the cyclopentene phosphate **5.3** to outcompete PRPP suggests that the altered geometry around the reaction centre is an important aspect of the inhibitor which may not be captured enough by the iminoribitol based analogue. An explanation could be the presence of a hydrogen at what would normally be the anomeric centre. This could introduce steric

constraints with the ATP substrate known to bind first at the active site. As a result this may place the ribose ring in a less optimal position for binding and overall result in no observable binding under competitive conditions. Extending the iminoribitol towards the ATP binding site by introducing features of the adenine ring may be necessary to overcome these constraints.

Chapter 6

Fragment based lead discovery

6.1 Fragment based lead discovery

Fragment based lead discovery (FBLD) is a method proven to be successful in drug design and development, with the first fragment derived drug approved by the United States food and drug association (U.S. FDA) in 2011, and with another 25 drugs in clinical trials.[160,161] What is of particular interest is that FBLD has been able to provide lead compounds where traditional methods such as high throughput screening (HTS) have been unsuccessful. Probably the most successful example of this is the development of beta-secretase 1 (BACE1) inhibitors, with four drugs in clinical trials, all of which were derived from fragments.[160,162,163] The central idea of fragment based lead discovery is to screen a target against low-molecular-weight compounds (≤300 mw) to identify relatively poor inhibitors. Although the fragments only bind with low affinity (K_d in micromolar to millimollar range), their small size allows exploration of different parts of the target enzyme active site with different fragments making the hits highly ligand efficient (LE).[164,165] Due to the smaller starting point there is more space to further optimise the structure by traditional medicinal chemistry with more direct control over the properties of the derived drugs. When multiple hits are found merging and linking these specific binders can rapidly enhance potency into lead-like compounds.

The definition of a fragment varies in the literature, but generally compounds with a maximum of 14 to 20 (150 to 350 mw) non hydrogen atoms are included.[166–168] As with Lipinski's rule of five to evaluate the "druglikeness" of a compound, a similar set of criteria has been established for FBLD called the rule of three.[169] The rule of three provides a guideline for fragments with a molecular weight cut off ≤300 Da, the number of hydrogen donors ≤3, the number of hydrogen acceptors ≤3 and a clogP of ≤3. In addition unwanted functional groups associated with toxicity, or functionalities that lead to false positives (aldehydes, disulphide bonds, etc.) are excluded from the compound libraries.[166] The next step then is very much dependent on the target (or targets) and chosen method of screening. Physical methods for screening these low affinity binders such as nuclear magnetic resonance (NMR) and thermal shift measurements only provide a medium throughput set up, therefore screening large libraries of several thousand compounds is not ideal.[170] To ensure that the full chemical space is included in a library a diversity screen is sometimes used. A diversity screen looks at the chemical similarity of compounds within a

set based on functionalities, such as hydrogen donor, hydrogen acceptor, or hydrophobic group, and the distance between the functionalities within a molecule. Compounds with the same functionality profile are excluded leading to a diverse small library.[171] Virtual screening provides another platform which allows for a more target focussed subset of compounds to be selected from larger libraries and can substantially reduce the amount of compounds that are physically screened. Virtual screening uses computational algorithms to predict the binding mode of compounds to a predefined active site.[172,173] Scoring function, mathematical methods to predict the strength of the non-covalent interactions of a ligand and receptor, can then be used to rank the individual compounds according to the quality of the interactions. The combination of predicted binding orientation and binding affinity is commonly referred to as docking. With the growing interest and success of FBLD substantial efforts have been made towards fragment docking and a number of successful examples are available.[174–176] One of the big challenges generally in docking experiments is the correct scoring of the binding energies of the different binding poses. In fragment docking these challenges are amplified in the correct scoring of small differences in binding energy for different binding poses of small molecules as well as distinguishing quality of binding between different fragments.[177]

After a fragment library or subset has been obtained the next challenge is screening the fragments for binding. The Ligands bind with relatively low affinity, it is therefore essential to confirm binding through a variety of techniques. The most common techniques used are NMR, surface plasma resonance (SPR), thermal shift assays and X-ray crystallography.[161,178,179] Of these techniques X-ray crystallography has of particular advantage that structural information can be obtained to help make decisions for further development of fragments into more lead like compounds easier. When a fragment shows a positive result in a combination of techniques this can be considered as a positive hit and a starting point for subsequent structure activity relationship (SAR) studies.

The ATP-PRTase active site was considered an ideal target for this approach as it has a large and solvent exposed active site providing a platform for multiple fragments to bind. This chapter reports on the virtual screening of a fragment library and ligand screening effort in an attempt to find novel scaffolds to inhibit ATP-PRTase from *M. tuberculosis*. Glide XP part

of the Schrödinger software suite was used for docking and the process of receptor preparation, docking procedure and selection of potential hits will be discussed.

6.2 Fragment library docking

6.2.1 Docking procedure

The general docking process can be divided in three parts.[180]

1. Receptor grid generation: The computational costs of evaluating a target are minimised in Glide by employing a grid based approach. A grid is a three dimensional description of the binding site which is generated by pre-calculating the interaction-potentials for each point on the grid. The grid only has to be calculated once which significantly speeds up the subsequent evaluation of docking poses.[181]
2. The docking methodology employed by Glide uses a series of filters also referred to as a "docking funnel" approach. The first step is the evaluation of the initial ligand conformations, pre-generated minima in the ligand torsion-angle space in the absence of the receptor, across the entire phase space. The poses selected by the initial screen are then further optimised in presence of the receptor active site and subjected to more exhaustive conformational searches.
3. The final step is the scoring function. Glide scoring is the empirically based ChemScore function of Eldrige *et al.* which describes contributions and quality of hydrophobic interactions, ligand receptor H-bonding interactions, and Coulomb interactions.[181,182] Steric clashes are panelised. Glide employs two forms of GlideScore: (i) GlideScore (SP), used by standard precision Glide docking; (ii) GlideScore XP, used by extra-precision Glide docking. The SP scoring function is a "softer" more forgiving function and is aimed to minimize false negatives. In contrast the XP gives larger penalties for poses that violate physical chemistry principles and aims to minimize false positives and can be considered in studies where limited numbers of compounds will be screened physically.

6.2.2 Receptor preparation

The receptor preparation and the virtual screen were carried out by Dr. W. Jiao. The crystal structure of *M. tuberculosis* ATP-PRTase in complex with the inhibitor AMP and allosteric inhibitor histidine was used as the receptor for docking (PDB: 1NH8).[47] The receptor

preparation was done following the standard procedure used in protein preparation wizard. In short, all the water molecules and sulphate ions present in the original structure were removed except for the active site ligand AMP, this was kept in the structure to help define the centre of the grid file later. The missing hydrogens from the experimental structure were placed back in and the H-bonds optimised. After this the structure was minimized. The grid file for virtual screening was generated before the virtual screening workflow. Using the prepared structure the centre of the grid is defined by centroid of residues within 8 Å of Tyr116 and, the grid size was set to "dock ligands with length ≤ 16 Å".

6.2.3 Preparation of the library

The fragment library from the Zinc database of commercially available compounds was used for docking.[183] The fragment library is a collection of over 500,000 compounds with calculated logP values of <= 3.5, molecular weight <= 250 Da, the number of rotational bonds <= 5 and unwanted functionalities are removed.[183] This set was prepared for docking applying Ligprep, part of the Schrödinger program package. The fragment library was searched for undefined chiral centres, which were removed, and the ligand protonation state was set to be consistent with to pH 8.0 ± 0.5.

6.2.4 Docking results and compound selection

The compounds were ranked by GlideXP docking score and the top 500 poses were visually inspected. Compounds selections were based on the following criteria:

- Interactions with key residues in the active site: Ligands that were predicted to form positive interactions with residues involved in substrate binding were chosen preferentially.
- Complementary fit to the binding site: Ligands whose van der Waals radii structure gave a tight fit with the active sites geometry, e.g. filling up of buried pockets and complementary electrostatic potential.
- Good geometry of ligand-receptor H-bonds: Interactions between a H-bond donor atom X-H and an acceptor atom :Y with a free pair of electrons were regarded as H-bonds if the distance between X and Y was between 1.8 and 3.0 Å, and the X-H-Y angle was larger than 120^{o}. Ligands with H-bond angles closer to 180^{o} were chosen preferentially.

- Absence of steric clashes: Small clashes between ligand and receptor as well as interamolecular clashes were tolerated, however clashes between ligand and key residues, multiple small clashed, or any bad clashes were penalised and led to rejection of the ligand.
- Potential for further development: Predicted binding poses that show potential for elaboration to other part of the active site where prioritised.

The active site of *M. tuberculosis* can be divided in to two sub-sites namely the ATP binding site and the PRPP binding site (Section 1.10).[38] The key residues are the charged phosphate binding loop (Gly 157- Arg 160) and Glu 141 and Asp 154 which are involved in binding the 2,3-OH groups of the ribose ring. At the time of this research there was no structure available with ATP bound to the active site of an ATP-PRTase enzyme. Sacchettini *et al.* used information from the *E. coli* ATP-PRTase structure with PRATP partially bound (PDB 1Q1K) to dock the PRATP product into the active site of *M. tuberculosis* ATP-PRTase (1NH8). Information on the ATP binding site was taken from this study. This work suggested that Asp 70 interacts with the ribose 2,3-hydroxyl groups and Tyr 116 and Lys 51 formed interaction with the triphosphate part of ATP. These residues were considered essential to ligand binding and fragments interacting with these portions were chosen preferentially. Selections were further based on the position of the fragments across the active site to ensure an even distribution and to increase the chance of finding complementary small ligands. Based on these criteria a shortlist of approximately 40 compounds was selected which were checked for availability, this gave a final short list of 16 compounds that were commercially available in sufficient quantity (Figure 6.1).

The predicted poses of the purchased fragments by docking are shown below, yellow dashed line shows the predicted H-bonding interactions. Non polar hydrogen atoms are omitted from the receptor for clarity and some of the key amino acids are labelled.

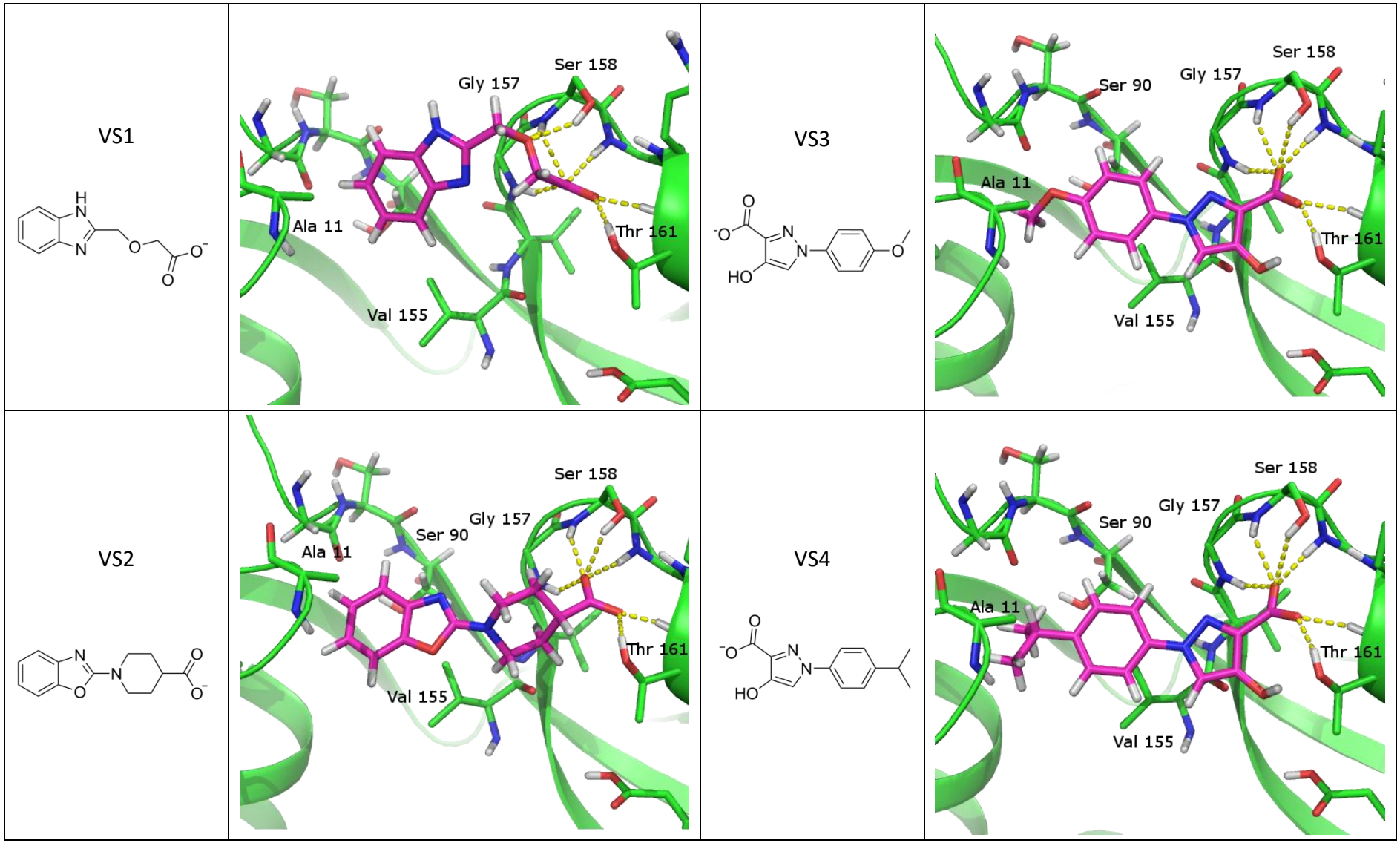


Figure 6.1. Structures and docking poses of compounds selected and purchased based on the virtual screen of the Zinc fragment library. To the active site of ATP-PRTase from M. tuberculosis.

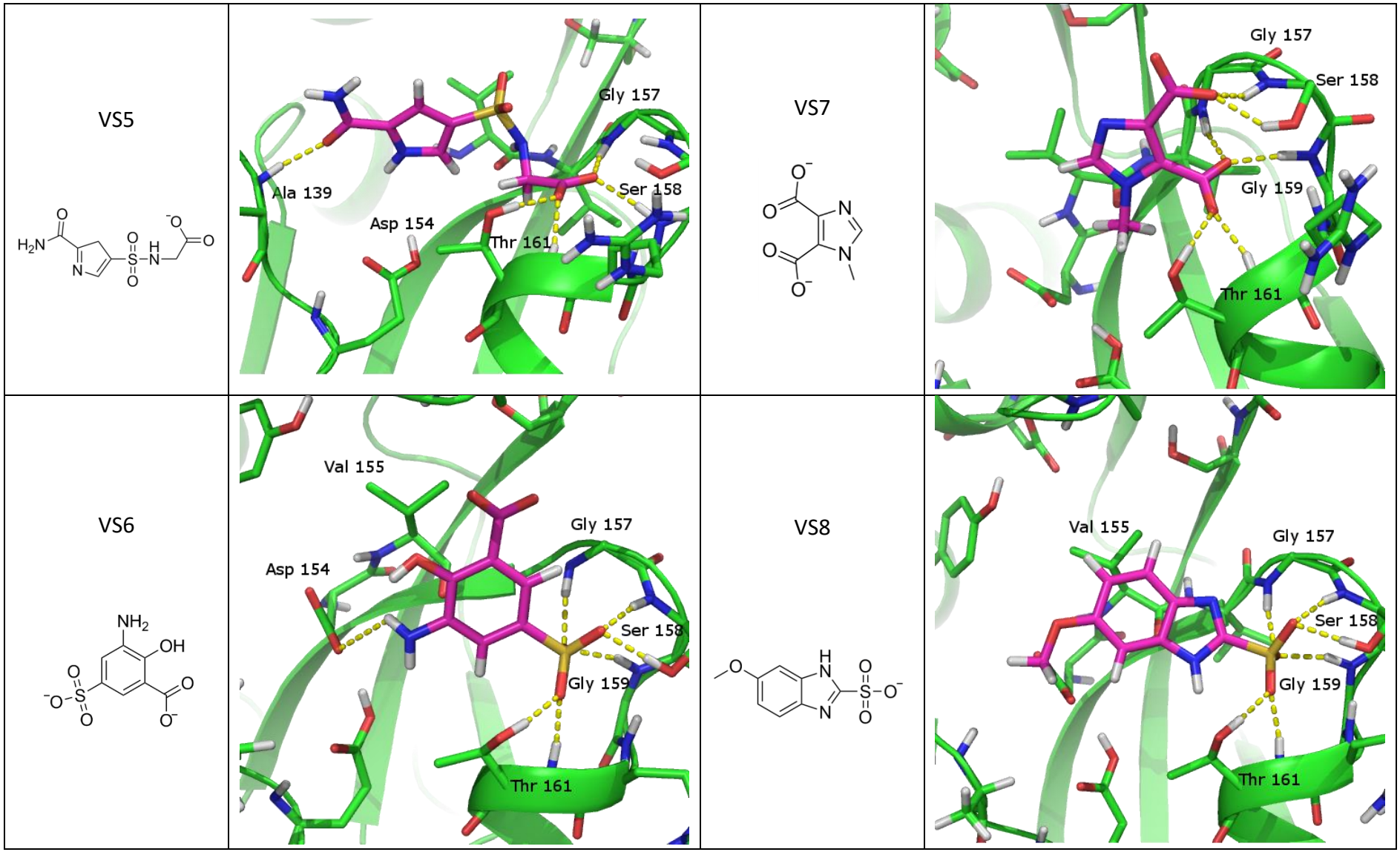


Figure 6. 1. (continued)

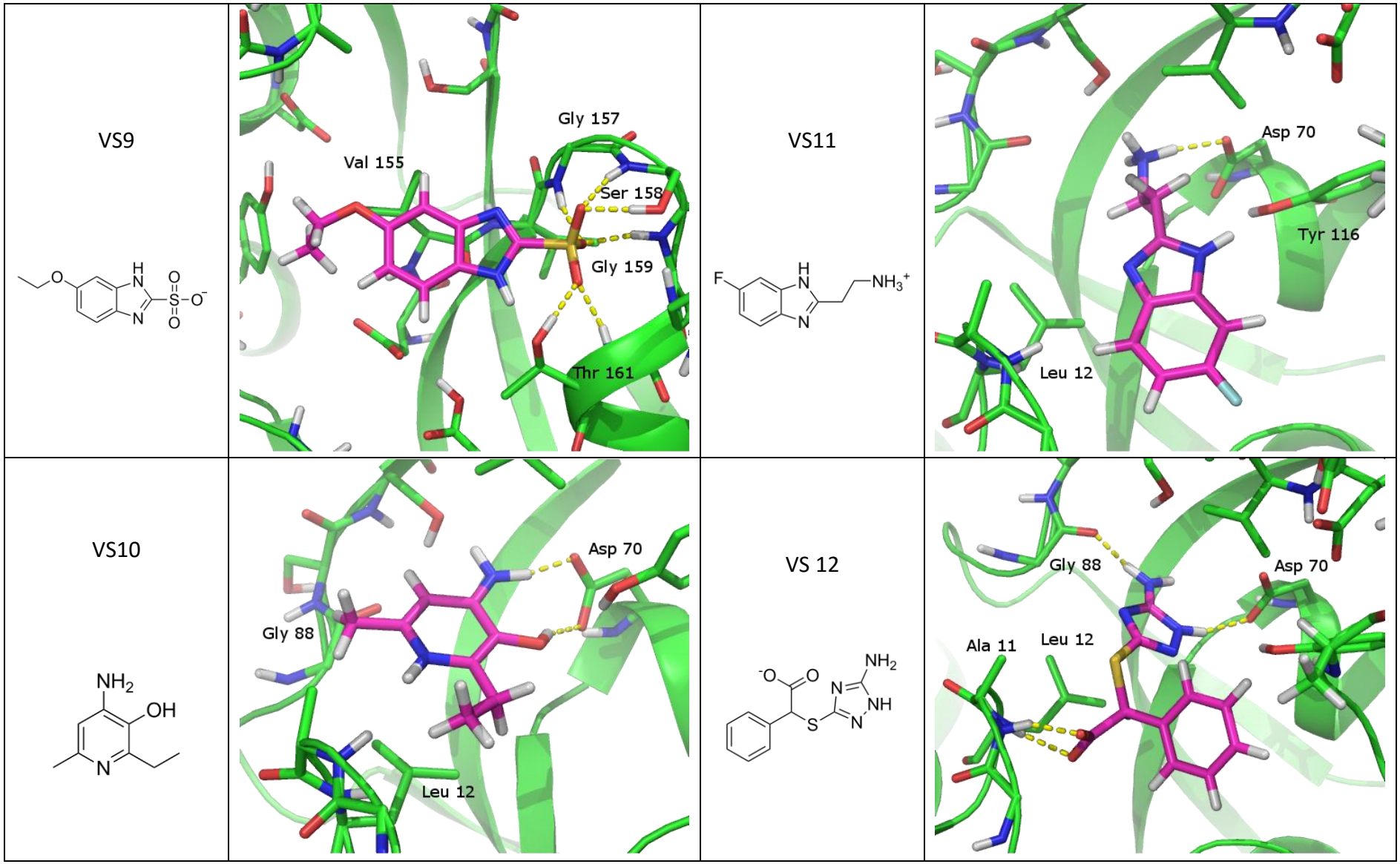


Figure 6.1 (continued)

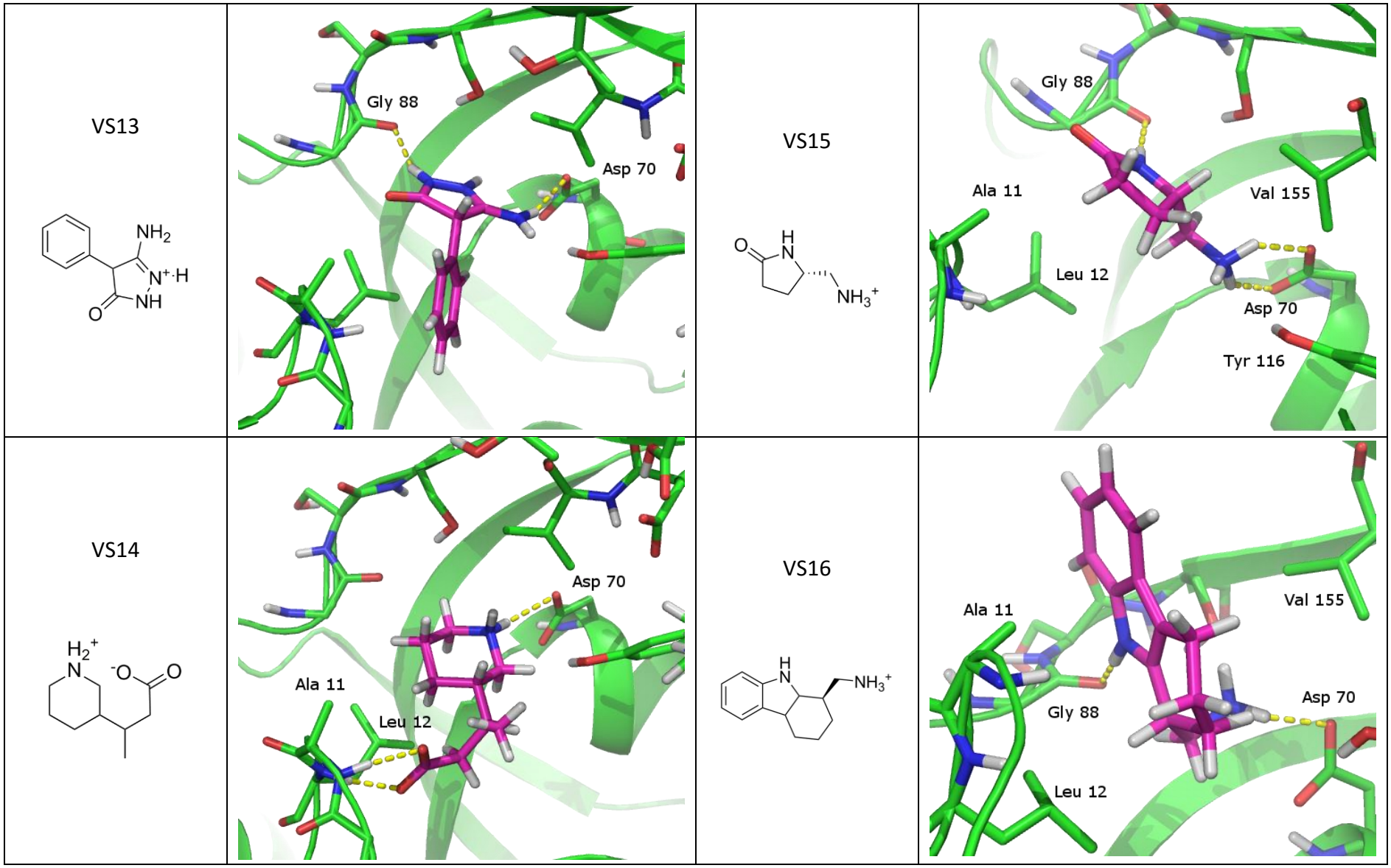


Figure 6.2.1. (continued)

The selected fragments can be divided into three groups based on the predicted position of binding. VS1-VS4 generally interact with the phosphate binding loop through a charged head group and further reach towards a hydrophobic envelope part of the predicted ATP binding site. VS5-VS9 also interact with the charged phosphate binding loop but the side groups bridge towards Glu141 and Asp154 and mostly occupy the PRPP binding site. VS10-VS16 are found at the ATP binding site in the pocket buried deep into the active site of ATP-PRTase, the side chains mostly sit in a hydrophobic groove.

A more detailed evaluation of the purchased fragments (Figure 6.1) is given below.

- VS1 is docked with a rich H-bonding network to the phosphate binding loop. Interactions are made with both the carboxylate head group, as well as the ether linker. The hydrophobic benzimidazole ring has a complementary fit to the active site were the purine base is expected to bind.
- The predicted docking pose for VS2 forms multiple H-bonding interactions with a good geometry to the carboxylate head group. The six membered ring has a complementary fit to the active site and the benzoxazole interacting with the adenine binding envelope between Ala11 and Val155.
- VS3 and VS4 share the same charged head group and were chosen for their fit to the PRPP binding site. The substituted aromatic rings were chosen to investigate the effect of a slightly different substitution pattern.
- VS5 was the only available fragment that reached all the way across the PRPP binding site forming both H-bonding interactions at the phosphate binding loop and the backbone oxygen of Ala139. The pyrrole nitrogen could provide an easy way to further functionalise this molecule.
- VS6 The sulphate group resembles the natural substrate's 5-phosphate group and gave a complementary fit to this site. The substitution pattern around the ring could allow for rapid expansion of this ligand into other parts of the active site.
- VS7 the size and multitude of H-bonding around the two carboxyl groups made this an interesting ligand.
- VS8 and VS9 show the same complementary fit of the sulphate to the phosphate binding loop as discussed for VS6 and the close resemblance to the adenine ring led to the selection of these ligands.

- VS10 the amine and hydroxyl provide H-bonding with good geometry and the aromatic ring occupies the buried pocket without many bad interactions. This also has potential for development towards other part of the active site.
- VS11 shows an H-bond with Asp70 and a good complementary fit to the hydrophobic binding pocket towards Lys51.
- VS12 the triazole gave a good fit to the buried pocket as well as additional H-bonding to the backbone amide of Gly88. In addition the carboxylate gave H-bonds with good geometry to the backbone amide of Ala11 and Leu12.
- VS13 is a similar to VS12, again the pyrazole ring complements the buried pocket and the aromatic ring is accommodates a hydrophobic part of the active site.
- VS14 again shows a complementary fit to the active site and H-bonding with good geometry around the carboxylic acid. In addition a hydrogen bonding interaction between the secondary amine and Asp70 was predicted.
- VS15 the primary ammonium ion forms H-bond with good geometry with Asp 70 and the five membered ring has a complementary fit to the active site.
- VS16 this ligand was selected for the complementary fit to the ATP binding pocket.

6.3 Hit validation

Validation of fragment binding to a target of interest is one of the challenging aspects of fragment based lead discovery, not only do these fragments bind with relative low binding constants (K_d), there is the potential for nonspecific binding to either the surface of the protein and/or other site such as allosteric sites. To rule out false positives or false negatives ideally a fragment is shown to interact with a target through a number of techniques. In this study three techniques were attempted, namely; detection by nuclear magnetic resonance (NMR), differential scanning fluorimetry (DSF), and crystallography. It is important to note that for a part of this research the *M. tuberculosis* ATP-PRTase enzyme was not available in the laboratory. The *C. jejuni* wt ATP-PRTase and a core mutant of this enzyme, which lacks the regulatory domain (domain III), were available and some of the techniques were set up using these enzymes. The *C. jejuni* ATP-PRTase enzyme has variations around the active site such as two charged residues around the ATP binding site where the triphosphate part is expected to bind and an arginine instead of Ala11. The arginine sits over the adenine binding envelope. Despite these differences the *C. jejuni* ATP-PRTase enzymes were

considered a reasonable model system based on similar conserved residues around the rest of the active site. In addition small fragments have been shown to interact across targets previously with specificity towards a target built in to the inhibitor controlled in the SAR stage.[184]

6.3.1 Saturation transfer diffusion NMR

Saturation transfer diffusion (STD) nuclear magnetic resonance (NMR) relies on the ability of an enzyme to transfer excitation to interacting ligands. The longitudinal relaxation rate (R_1) of protons in proteins is dominated by cross relaxation. This allows for selective saturation of a single resonance of a protein and through rapid spread of the magnetisation over the protein (spin diffusion) the entire protein is saturated.[170] When a ligand is interacting reversibly with the protein, the ligand will also get saturated. Because the relaxation of small molecules is slower than that of the protein the amount of saturated ligand builds up, which can subsequently be detected. Because only ligands that interact with the protein are excited, these can be visualised specifically.[170,185]

This work was carried out in collaboration with Dr H. Wong and Prof. P. Tyler part of the Ferrier Research Institute. At the start of the project this technique was not established at the University of Canterbury or the Ferrier research institute and work towards establishing this technique was carried out first. This work was carried out on the *C. jejuni* wt ATP-PRTase enzyme. NMR techniques require longer experiment times so first the stability of the enzyme was tested over a six day time period with the specific activity as a measure of degradation (appendix I). The enzyme was found sufficiently stable for up to two days under conditions similar to that of the NMR experiment. The technique was set up using the known inhibitor AMP in presence of two non-binders cyclohexylamine and tryptophan at 2 mM. After optimisation no signal was observed in absence of the enzyme (Appendix III) when the enzyme was added to the sample AMP showed stronger excitation then two non-binders. A small signal was still observed also for these compounds. The difference was considered sufficiently different to identify non binders (Appendix III). When the experiment was repeated with a ligand mixture of VS2, VS4 and VS15 two of the three ligands showed to interact (Figure 6.2). The ligands that gave a positive signal were VS2 and VS4 predicted to bind at the phosphate binding loop towards the adenine binding envelope. To test specificity of binding the competitive inhibitor AMP was added, which has been shown

crystallographically to bind across this part of the active site as well (Section 1.8.1.2). Known binders should compete for binding at the active site and reduce or completely remove signals from other ligands. AMP however was unable to reduce the signals arising from the tested ligands suggesting non-specific interactions. With the technique being particularly time intensive and the limited success this technique was abandoned and other screening techniques were considered to test the selected fragments for binding.

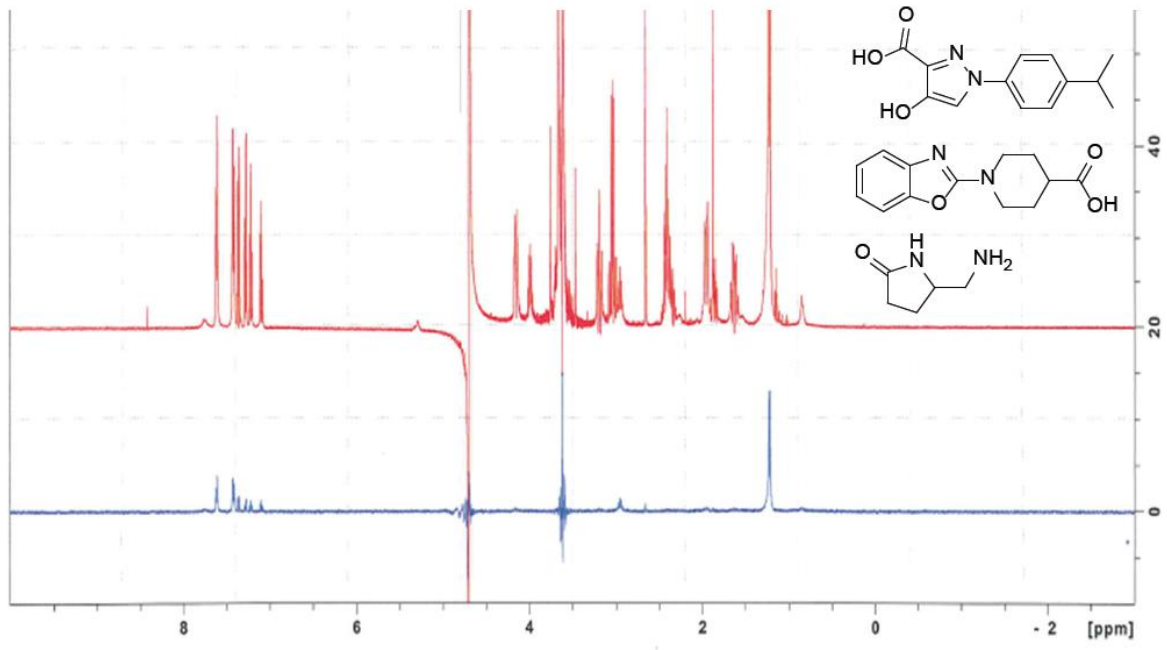


Figure 6.2. In blue is shown the STD-difference NMR spectrum with signals from VS2 and VS4. In red is a normal 1D-proton spectrum with water suppression of the same sample.

6.3.2 Differential scanning fluorimetry

Differential scanning fluorimetry (DSF) is a thermal shift assay that quantifies the melting temperature (T_m) of a protein. Interaction between a ligand and protein usually induces changes in protein thermal stability which modifies the midpoint denaturation temperature.[186] In a DSF experiments this shows as an increase or decrease in melting temperature and the extent of the stabilisation or destabilisation is an indication of the strength of binding. In drug discovery efforts generally a stabilising effect, which indicates enthalpically driven binding, is desirable for the initial starting point.[161,187–189] An enthalpic binding event is commonly associated with the formation of H-bonds and sometimes used as an indicator for specific molecular interactions with an active site.[188,190,191] Although DSF

is only a crude indicator for binding the low material cost and relative high throughput screening makes it a good initial screening technique. One of the challenges in DSF is that small changes in T_m, 0.5-2 °C, may be hidden by the melting profile of a protein through a slow transition or multiple small denaturing events.[192] Unfavourable melting profiles were observed for both *M. tuberculosis* and *C. jejuni* ATP-PRTase wt enzyme (Figure 6.3, **A** and **B**). AMP, a known inhibitor of the ATP-PRTase enzymes, gave no observable change in melting temperature for the *M. tuberculosis* ATP-PRTase enzyme (Figure 6.3, **A**). With no change in melting temperature observed for the natural substrate this technique was considered unsuitable for screening the ligands against *M. tuberculosis* ATP-PRTase. *C. jejuni* ATP-PRTase gave a suboptimal baseline however a difference for AMP binding was observed (Figure 6.3, **B**). The functional *C. jejuni* ATP-PRTase core mutant displayed a much better T_m profile for both the unliganded and in presence of AMP (Figure 6.3, **C**). Based on the melting profile the core mutant was considered a better candidate for the DSF work. The fragments were screened again the *C. jejuni* ATP-PRTase core mutant at a concentration of 2 mM and using replicates of three (Table 6.1). The blank protein, only containing a small amount of DMSO, gave a melting temperature of 48.7 ± 0.2 °C ligand able to invoke a change ≥ ± 1.5 °C were considered.

The fragment screen showed no stabilisation of the target protein for any of the screened ligands. Three ligands showed destabilising effects to the protein by more than 1.5 °C with the strongest response from VS4, yielding a destabilising effect of -4.3 ± 0.2°C, and VS3 and VS13 -1.5 ± 0.2 °C and -2.1 ± 0.2 °C respectively. VS4 also showed a positive response in STD-NMR however specificity of binding could not be confirmed. The based on the NMR results the strong destabilising effect was attributed to nonspecific binding of this ligand to the target protein.

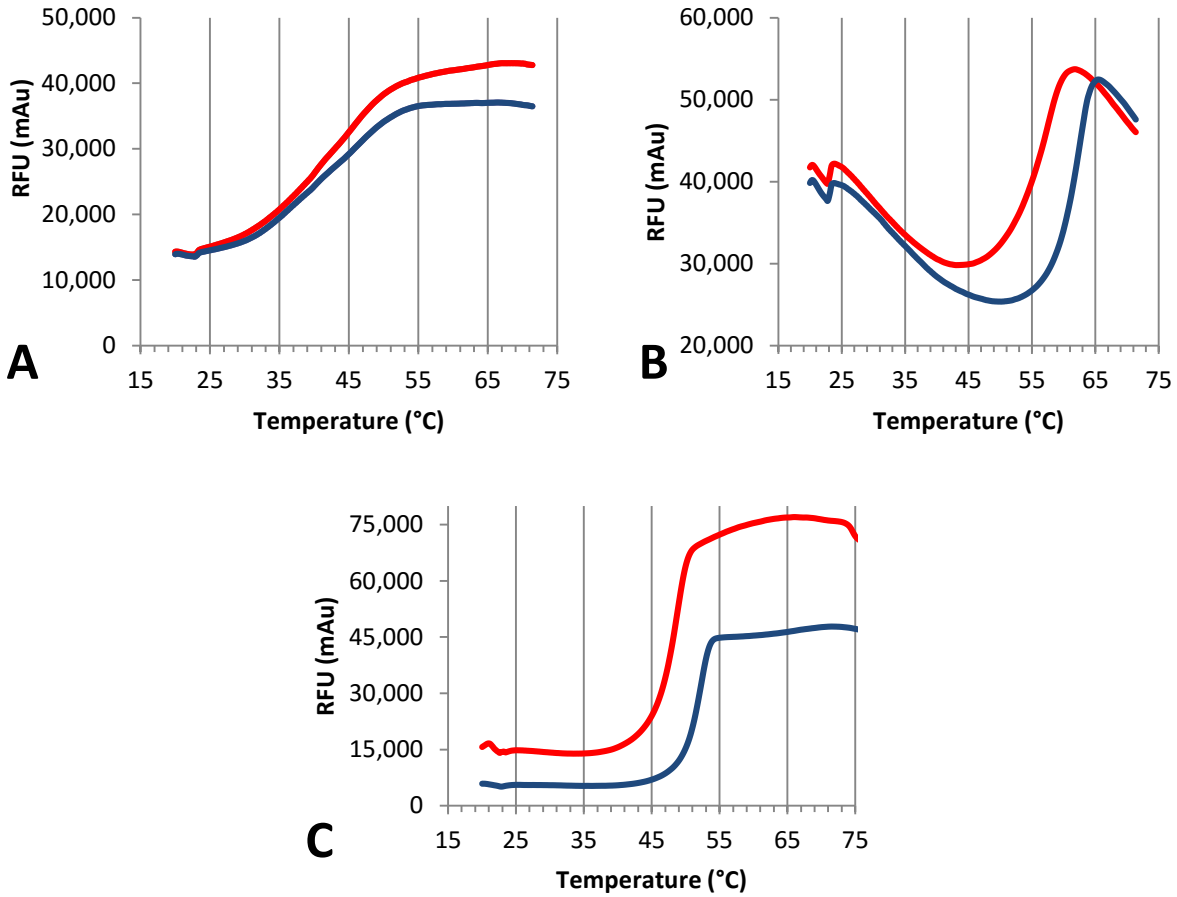
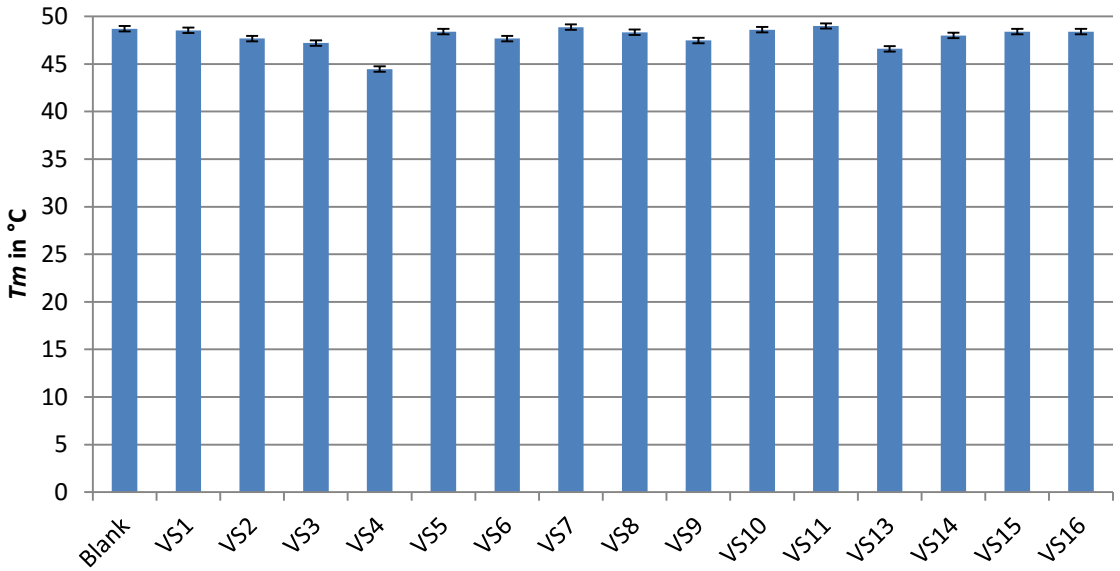


*Figure 6.3. The melting profile for M. tuberculosis ATP-PRTase(**A**), C. jejuni wt ATp-PRTase(**B**), and C. jejuni core mutant ATP-PRTase(**C**). Red line shows the melting curve for the enzyme in absence of ligands, blue line shows the change in melting temperature in presence of known competitive inhibitor AMP at 1 mM. Specific melting temperatures where found from the inflection point of the first derivative curve.*

Table 6.1. Results from the fragment screen by DSF for the C. jejuni ATP-PRTase core mutant. The enzyme melting temperature is given in °C. Ligand 8 reacted with the dye and was omitted. Error bars are the STD.



6.3.3 Crystallography

Conditions to grow apo crystals for *C. jejuni* wt ATP-PRTase enzyme were not available. The *C. jejuni* ATP-PRTase core mutant on the other hand was successfully crystallised in the apo form and the natural ligands ATP, PRPP, and PRATP successfully soaked into these crystals at good resolution 1.8-2.2 Å.[67] The *C. jejuni* ATP-PRTase core mutant was crystallised according to established procedures, 0.1 M sodium acetate, 0.01 M $ZnCl_2$, 7-10 % w/v PEG 6000, pH 5.0 protein concentration ≥10 mg/mL.[67] Crystals for *M. tuberculosis* were grown according to published conditions, 1.6 M $MgSO_4$, MES pH 6.5 at protein concentration ranging from 15 to 20 mg/mL.[47] Ligands were either co-crystallised at concentrations ranging from 5 to 50 mM depending on the solubility of the ligand or the apo crystal soaked in presence of the ligand at concentrations within the same range. Generally crystals that did not break up or dissolve during this treatment were looped, cryoprotected and flash frozen 30-120 min afterwards. Several datasets were collected for both enzymes however no ligand density was observed for any of the ligands.

6.4 Discussion

There were several factors that could contribute to the unsuccessful outcome of this project. Firstly, although docking of fragments has been successful and examples have been published, correctly docking and scoring of small molecule to a target still remains challenging. The incorrect docking tends to arise from incorrect scoring of the generated binding poses which in turn arises from the small difference in binding energy for different binding poses of the small molecules.[177] The incorrect scoring also affects discrimination between different fragments and makes distinguishing actual binders harder. In addition to test the ability of a docking program to correctly dock a ligand the native substrate and/or inhibitors could be used to ensure the ability of the program to correctly model the binding poses. Furthermore only a small subset of the high ranking ligands has been purchased for testing, the information from the initial screen could be used to feed back into finding alternative fragments.

Effective screening by DSF of the fragment library was challenging for *M. tuberculosis* ATP-PRTase with the melting profile hiding stabilisation for the natural ligands making the technique unsuitable for this enzyme. The *C. jejuni* ATP-PRTase core mutant allowed for screening of the fragment library, however no stabilisation was observed for any of the

selected compounds. With the library tailored for the *M. tuberculosis* ATP-PRTase enzyme and the differences at the active site, such as an Arg16 in place of an alanine in *M. tuberculosis* ATP-PRTase and the later revealed difference in binding mode of ATP, this result is maybe not surprising. Three ligands were found to destabilise the *C. jejuni* ATP-PRTase core mutant, however their specificity of binding could not be established through crystallographic efforts. Crystallisation efforts were made for both enzymes through co-crystallisation and soaking experiment of all ligands however this did not yield any observable ligand density in any of the structures.

Chapter 7

Conclusions and future directions

The work discussed in this thesis contributes to the knowledge of the structure and function and inhibition of the ATP-PRTase enzyme. All the initial research aims were achieved, with two crystal structures of *C. jejuni* ATP-PRTase providing further insight into substrate binding and how this enzyme may function as a hexameric species. Based on the structures a structural basis for allosteric inhibition has been proposed. The reaction mechanism was successfully investigated through KIE studies and transition state analysis. Furthermore, inhibition of the ATP-PRTase enzyme through non-natural ligands was achieved. These findings are expanded below.

7.1 Structural basis for the inhibition of *C. jejuni* ATP-PRTase

A crystal structure for the *C. jejuni* ATP-PRTase enzyme was obtained with the natural substrate ATP bound at the active site. This is the first time clear density for this ligand has been reported at the active site of the ATP-PRTase enzyme. Previously a dynamic equilibrium of oligomeric states with an active dimer and the inhibited hexamer had been suggested.[193] However, solution studies performed prior to this work did not show a dimeric species even at low enzyme concentration, making a dynamic equilibrium unlikely and supporting a functional hexamer.[66] Analysis of the interface of the active ATP-bound crystal structure further supports this finding with a more significant contribution of the trimer interface contributing to the hexamer. The hexameric arrangement places the active site cavity on the inside of the hexamer, raising the question of how substrates and products enter and leave the overall structure. The active ATP bound form of the enzyme has a large entry at the trimer interface, and a small entrance can be observed at the dimer interface which may provide direct access to the active site. A second structure with histidine and AMP bound to *C. jejuni* ATP-PRTase was also obtained. Comparison of the AMP binding mode to the ATP-bound structure supports the competitive nature of AMP towards both substrates ATP and PRPP with the competitive inhibitor binding across both sites. These findings were consistent with the previously proposed mechanism. The inhibitor-bound structure provided further insight into the origin of the reduced activity of the protein in the inhibited form and an alternative mechanism of inhibition has been proposed which does not require a dimer hexamer equilibrium. The allosterically inhibited structure displays a more rigid and narrow structure then the ATP bound form with a more extensive trimer

interface. This results from a regulatory twisting motion which brings the monomeric units closer together, reducing the size of the entrance at the trimer interface and limiting the access into the hexamer interior. Histidine binding further stabilises the hexamer, demonstrated by a thermal shift assay which resulted in an increase in thermal stability upon binding of the allosteric ligand.[66] The reduced dynamics are likely to be partly responsible for loss of activity of the inhibited structure.

7.1.1 ATP binding to ATP-PRTase

In addition to the ATP bound structure for the *C. jejuni* ATP-PRTase enzyme, a crystal structure with ATP bound at the active site of *M. tuberculosis* ATP-PRTase was also obtained. With the ATP binding site only partially conserved in these two ATP-PRTases, the two structures provide insight into how these differences impact the binding mode. The binding of the adenosine part was found to be very similar regardless of the variation in amino acids around the adenine base. In the *C. jejuni* ATP-PRTase the adenine base binds in a hydrophobic region of the active site with an arginine residue (Arg16) sitting over the top forming a π-π stacking interaction. Although directly involved in binding, this residue is not conserved in *M. tuberculosis* ATP-PRTase, where an alanine (Ala11) is found in this position. The triphosphate group of ATP for *C. jejuni* ATP-PRTase is held in place through strong interactions with a double aspartate motif through a divalent magnesium cation. This motif is not conserved among ATP-PRTase enzymes, and for the *M. tuberculosis* ATP-PRTase a proline is found at this position. Binding of the triphosphate is still associated with Mg^{2+}, however fewer interactions occur with the active site directly. Coordination of Mg^{2+} to the triphosphate places the cation in the centre of the active site. An additional Mg^{2+} cation is found at the dimer interface where it is coordinated to the γ-phosphate of two ATP ligands. These structures provide the first example of ATP binding at the active site of ATP-PRTase enzymes.

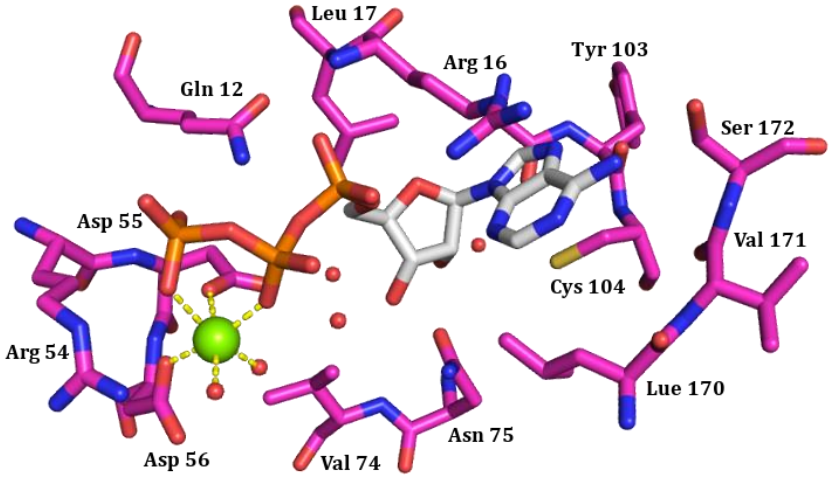


Figure 7.1. ATP binding at the active site of C. jejuni ATP-PRTase with the interactions around Mg^{2+} indicated with yellow dashed lines. Main chain in purple, oxygen in red, nitrogen in blue, sulphur in yellow and phosphate in orange.

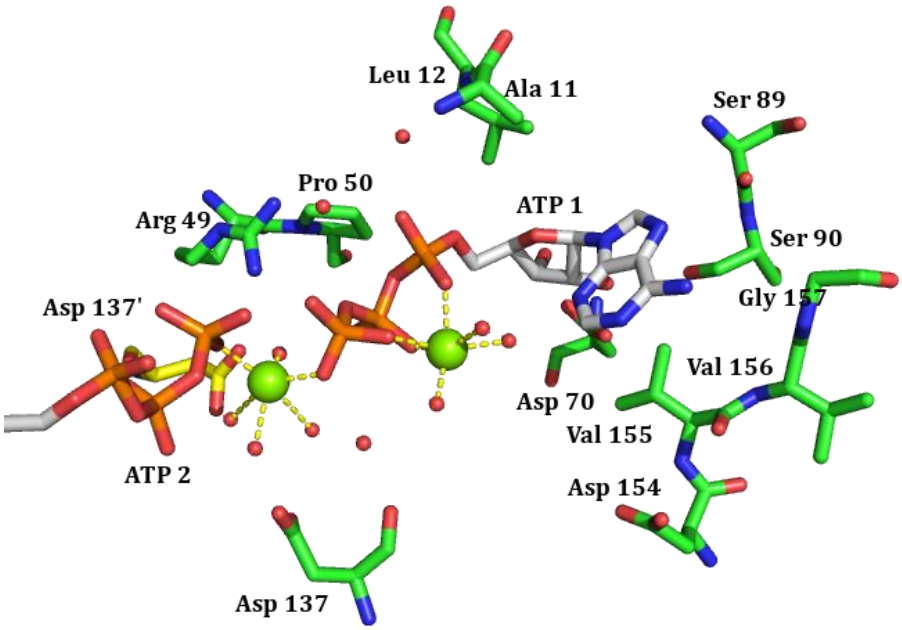


Figure 7.2. ATP binding at the active site of M. tuberculosis ATP-PRTase with the interactions around Mg^{2+} indicated with yellow dashed lines. Main chain in purple, oxygen in red, nitrogen in blue, sulphur in yellow and phosphate in orange.

7.2 The transition state of ATP-PRTase enzymes

A kinetic isotope effect guided transition state analysis was conducted for both the long and short form of ATP-PRTase. For this study, methods for the chemo-enzymatic synthesis and purification of isotopically labelled PRATP, and the conditions for the measurement of intrinsic KIEs through internal competition were established. The reaction was studied in the reverse direction which allowed commitment to catalysis to be overcome through the use of

the pyrophosphate alternative phosphonoacetic acid (PA), lowering commitment values to within experimental uncertainty and the measured KIEs were thus intrinsic. The primary ^{14}C and α-secondary ^{3}H KIEs were obtained for *M. tuberculosis*, *C. jejuni* and *L. lactis* ATP-PRTase, showing small primary KIEs of 1.028-1.031. This supports a dissociative reaction mechanism for the studied enzymes, further reinforced by the large α-secondary ^{3}H KIEs of 1.147-1.250. The involvement of ATP was established through ^{15}N isotope effects at the N1 and N6 position. Transition state models were generated with the experimental KIE values forming the boundary conditions for the structures. The transition state was located with density functional theory at the B3LYP/6-31^{+}G(d,p) level, and through several iterations a close match between the model and experimental values was found. The transition state models supported a dissociated reaction mechanism implied by both the primary and α-secondary KIEs. The ATP-PRTase enzymes for *M. tuberculosis*, *C. jejuni* (long forms) and *L. lactis* (short form) were all found to go through dissociative $D_N{*}A_N^{\ddagger}$ reaction mechanism with partial bonding to the PA nucleophile at the transition state and a well-developed carbocation (Figure 7.3).

The modelled nitrogen isotope effects support a fully dissociated adenine basis for all enzymes, however variations in the involvement of the adenine and tautomeric states were predicted. A large primary ^{15}N isotope effect of 1.024 ± 0.001 for *M. tuberculosis* ATP-PRTase supports a fully dissociated base as either the neutral N7 or N3 tautomer base, with the N3 tautomer considered more likely based on a conserved water molecule. For *L. lactis* ATP-PRTase a primary ^{15}N of similar magnitude to the *M. tuberculosis* enzyme was observed with a KIE of 1.028 ± 0.001. For this enzyme, the N7 protonated state might also be possible with the presence of Glu158 in close proximity of the adenine binding site. *C. jejuni* ATP-PRTase gave a ^{15}N KIE near unity at 0.995 ± 0.003. A primary nitrogen KIE at unity has not been observed previously for a nitrogenous base. A KIE near unity suggests that either no change in vibrational freedom takes place or that the bond changes are compensated for at the transition state. The difference may arise from a small difference in the overall reaction coordinates, with the adenine dissociation step for *C. jejuni* ATP-PRTase further separated from the transition state than it is in the *M. tuberculosis* and *L. lactis* ATP-PRTase catalysed reactions.

The involvement of the incoming nucleophile, phosphonoacetic acid, was still observed at the transition state with C1'-O^{PA} atom separation ranging from 2.50 Å for the *L. lactis* ATP-PRTase enzyme to 2.65 Å and 2.75 Å for *C. jejuni* and *M. tuberculosis* ATP-PRTase enzymes respectively. The model further revealed the importance of the catalytically essential Mg^{2+} cation in both anchoring the ribose-5-phosphate and nucleophile in place for effective catalysis. Furthermore, the Mg^{2+} helps stabilise the charged nucleophile at the transition state. The positively charged anomeric centre at the transition state was predicted to be stabilised by the ribose ring oxygen resulting in a flattened ring and a less negatively charged ribose ring oxygen, typical for ribocations (Figure 7.3).

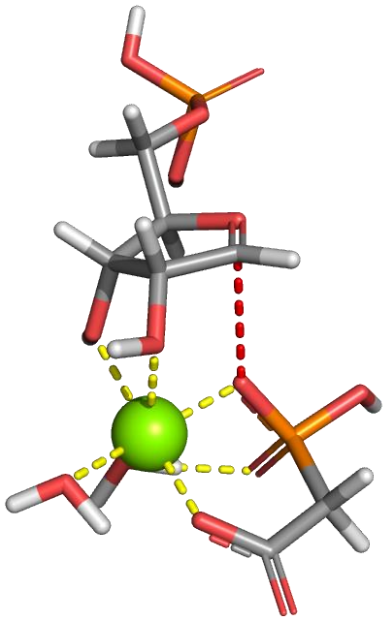


Figure 7.3. At the transition state a well-developed carbocation is formed with partial bonding to the incoming nucleophile, which agrees with a $D_N*A_N^{\ddagger}$ reaction mechanism.

Overall this study is the first detailed analysis of the transition state of adenosine triphosphate phosphoribosyltransferase with the transition state successfully modelled based on experimental intrinsic KIE values.

7.2.1 Future directions

The transition state models generated in these studies were all calculated *in vacuo* and although this has been used throughout literature to accurately model transition state structures, it still raises questions about how interactions from the active site might influence the transition state. A quantum mechanics/molecular mechanics (QM/MM)

approach may provide further insight into the role of the active site at the transition state. In particular the interactions of the active site to the adenine base may provide insight into the variations observed around the ^{15}N isotope effects, which are difficult to infer from *in vacuo* calculations alone.

The active sites of *M. tuberculosis*, *C. jejuni* and *L. lactis* ATP-PRTase all vary to some degree particularly around the binding site of the adenine base. Two of the enzymes bind the adenine part between the guanidinium group of arginine and hydrophobic residues. Although clearly involved in binding, this arginine residue is not conserved. In *M. tuberculosis* ATP-PRTase for example an alanine is found at this position. Presence of an alanine does not alter the binding mode around this part, how the difference in interactions influences the transition state is hard to predict. In addition for *C. jejuni* ATP-PRTase, a serine is found in close proximity to the adenine amino group and this residue is not observed for the other two enzymes. Mutation of some of the residues involved in adenine binding such as Arg16 to an alanine in *C. jejuni* ATP-PRTase, or Ala11 to an arginine in *M. tuberculosis* ATP-PRTase, and remeasuring the ^{15}N isotope effects could provide further insight into the importance of these interactions at the transition state.

7.3 Fragment based lead discovery

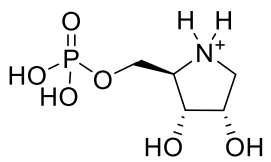
A fragment based lead discovery technique was explored for the *M. tuberculosis* ATP-PRTase enzyme through fragment docking, followed by screening the selected fragments for binding. Docking as the first screening technique allowed for a smaller set of compounds to be selected for physical screening. A challenge around docking fragments is that scoring the weaker interactions can be difficult. A library of over 500,000 fragments was docked into the active site of the target enzyme and a selection of compounds acquired. STD-NMR was chosen as the first screening technique; however the challenges around setting up the technique meant that the method was abandoned. An alternative screening technique was therefore explored. DSF provided a platform to screen fragments in a medium throughput set up. Here, unfavourable melting profiles made detection of ligand binding for the target protein *M. tuberculosis* ATP-PRTase challenging and the alternative *C. jejuni* ATP-PRTase core mutant was used for screening. This was unable to show any favourable binding from the selected fragments. The differences in active site residues particularly around the ATP binding may have influenced these results. Overall the uncertainties around the ligand

docking in combination with the challenge of screening these small molecules for binding to a target enzyme led to an unsatisfactory outcome of this particular part of the project.

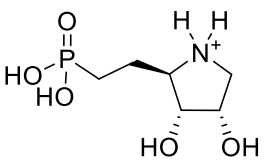
Although the FBLD technique presented was unsuccessful, the large solvent exposed active site the ATP-PRTase enzyme should provide a good target for FBLD. Screening of commercially available libraries has successfully been used to identify binding fragments and may form a better starting point for the screening regimen.[194–196] In addition, optimisation of the DSF conditions, such as alternative assay buffers, could improve the melting profile and allow measurement of the fragments to the *M. tuberculosis* ATP-PRTase enzyme to be made directly.

7.4 Synthesis and inhibition of ATP-PRTase through non-natural inhibition

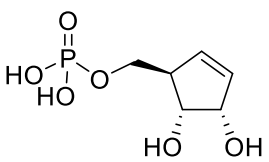
Three transition state analogues were designed, synthesised and tested for inhibition for inhibition for *M. tuberculosis* and *C. jejuni* ATP-PRTase. Although the transition state of ATP-PRTase was unknown at the time, based on related enzymes a dissociative reaction mechanism was expected prior to the design of the transition state analogues. Three intermediates were proposed, based on the carbocation intermediate. One analogue, the cyclopentene analogue, was designed to represent the altered geometry around the anomeric centre of the ribocation species. Two analogues, the iminoribitol phosphate and the iminoribostol phosphonate, were based around the charge change at the transition state (Figure 7.4). The cyclopentene analogue was shown to be a competitive inhibitor for the *M. tuberculosis* and *C. jejuni* ATP-PRTase enzymes. This result illustrates that even if small parts of the transition state are captured, inhibition may be observed. The iminoribitol based analogues were less successful, with no inhibition observed for either of these analogues under the conditions tested. The iminoribitol analogue does not capture the altered geometry around the anomeric centre. The presence of a hydrogen at this position may give rise to small steric clashes, making the analogue less favourable for binding.



5.1



5.2



5.3

Figure 7.4. The three synthesised transition state analogues.

7.3.1 Future directions

The transition state analogues only represent a small part of the full length ligands that are able to bind at the active site of ATP-PRTase. Although relatively weak, the observed competitive inhibition of the cyclopentene analogue towards PRPP is encouraging. To enhance potency, it would be interesting to capture more of the ATP binding interactions. A good starting point could be the adenine ring system. The transition state analysis supports a fully dissociated adenine base at least 3.0 Å away from the anomeric centre of the ribose ring for both *C. jejuni* and *M. tuberculosis* ATP-PRTase. A methylene or ethylene bridge might therefore need to be included to allow for appropriate positioning of both the adenine ring and the ribocation portions of the inhibitor.

For the cyclopentene based analogues, substitution directly to the double bond will likely place the substituent more planar to the ribose ring, making direct linking of the adenine at the N1 less than ideal. A crystal structure with PRATP bound to *C. jejuni* ATP-PRTase core mutant shows that the C2 could be a good alternative linking site (Figures 7.5 and 7.6).[67] Another advantage of linking to the C2 position is that it overcomes the problem around the Dimroth rearrangement shown to affect N1 substituted adenine rings.[99]

Additional structural information through crystallography of the cyclopentene inhibitor bound at the active site would provide valuable insight on potential alternative ways to link the ATP structure to the cyclopentene analogue, and other improvements.

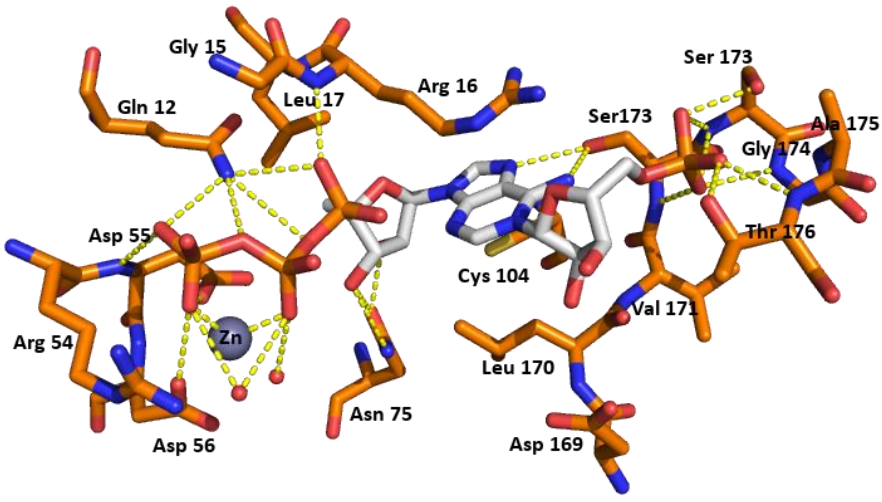


Figure 7.5. PRATP (white) bound at the active site of C. jejuni ATP-PRTase core mutant. Main chain in brown, oxygen in red, nitrogen in blue, sulphur in yellow and phosphate in orange. Polar interactions are highlighted by yellow dashed lines.

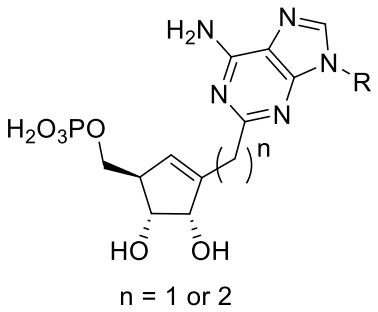


Figure 7.6. Example of an expanded cyclopentene-based transition state analogue.

Although the iminoribitol did not demonstrate any inhibition, based on previously reported success for these types of inhibitors, it would be worth extending this analogue. The addition of an aromatic functional group might place the iminoribitol ring in a better position for binding. Given the predicted bond distance of at least 3.0 Å between the N1 and the anomeric position it could be worth including a methylene bridge to better mimic this feature (Figure 7.7). Furthermore to overcome problems around the Dimroth rearrangement it might be advantages to replace the N1 with a carbon atom to provide a more stable analogue.

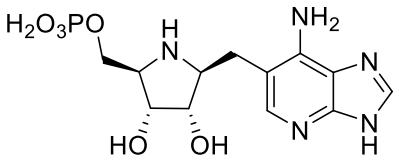


Figure 7.7. Proposed expansion of the iminoribitol based inhibitor. To ensure a stable intermediate the N1 might have to be replaced with a carbon atom.

Chapter 8

Methods and experimental procedures

8.1 General Experimental procedures

8.1.1 Reagents and solvents

Reagents were obtained from commercial suppliers and used without further purification unless stated otherwise. Before use, solvents were dried by passage through alumina columns in a solvent purification system custom-built by the Chemistry Department workshop according to published plans.[197]

8.1.2 Chromatography

Merck Kieselgel 60F254 pre-coated aluminium-backed plates were used for all thin layer chromatography (TLC). Visualisation of product was achieved by UV lamp at 254nm followed by staining with either basic permanganate solution or 10% H_2SO_4 in ethanol and heating in a stream of hot air. Amines were also visualised in acidified ethanolic ninhydrin. Flash chromatography was carried out on silica gel, 230-400mesh. Chromatographic petroleum ether and ethyl acetate were distilled on a rotary evaporator prior to use; analytical grade diethyl ether, toluene, methanol and ethanol were used without further purification.

8.1.3 pH measurements

A Denver Instruments UB-10 Ultra-Basic pH meter, with a standard probe was used for the determination of the pH of all buffers and other aqueous solutions. pH adjustments were made using hydrochloric acid, 1 M or 12 M, or sodium hydroxide, 1 M or 10 M. Volatile buffers were adjusted in pH by the addition of acetic acid as a 5% solution in Milli-Q or as the concentrated acid.

8.1.4 Mass spectrometry

The molecular mass of molecules and proteins were determined on by Mass spectroscopy on a Bruker maXis 3G UHF-TOF tandem mass spectrometer. Proteins were buffer exchanged into 10 mM $AcONH_4$ buffer and analysised at 1 mg/mL. Small molecules were dissolved in methanol, water, or methanol/dichloromethane mixtures depending on solubility.

8.1.5 Preparation of buffers and kinetic assay solutions

Buffers and kinetic assay solutions were prepared in water purified with a Milli-Q® intergral Water Purification system from Merck Millipore, this will be referred to Milli-Q throughout this thesis. After preparation and pH adjustments were made all buffers were filtered either through a 0.2 micron Durapore®membrane filter or 0.2 micron syringe filter. Assay solutions were pH adjusted and filtered through a 0.45 micron syringe filter.

8.1.6 Enzyme storage

Concentrated protein solutions (5-15 mg/mL depending on the enzyme) were dispensed into aliquots of 50 µL to 200 µL in volume and either flash frozen, by submerging in liquid nitrogen or packing in dry ice, or stored at 4 °C, immediately after purification depending on the intended use. Enzymes were kept on ice when in use.

8.1.7 determination of protein concentration

Protein concentrations were measured by absorbance at 280 nm on a Nanodrop® ND-1000 spectrometer blanked with the appropriate buffer. The protein concentration was calculated from the molar extinction coefficient or by 1% absorbance. Molar extinction coefficient values and percentile absorbance's were calculated from the protein sequence using the ProtParam tool on the ExPASy Proteomics server.

Enzyme	Molar extinction coefficient (M^{-1} cm^{-1})	Absorbance (1%) (=1 mg/mL)
C. jejuni ATP-PRTase wt	14440	4.29
C. jejuni ATP-PRTase core mutant	7450	3.31
L. lactis ATP-PRTase	26945	5.06
M. tuberculosis ATP-PRTase	19940	6.54

8.1.8 Protein concentration and buffer exchange

Proteins were concentrated by centrifugation in either a 5 ml or 20 mL concentrator with a 10,000 Da molecular weight cut off membrane. Buffer exchange was carried out by concentrating protein to a small volume followed by dilution with the desired buffer. This process was repeated at least 3 times.

8.1.9 Standard kinetics assay

The activity of ATP-PRTase enzymes was monitored at 290 nm following a modified methods by Morton and Parsons.[56] The kinetic assays were carried out on a Varian Cary 100

UV-visible spectrophotometer or a Varian Cary 300 UV-visible spectrophotometer. All reactions were carried out in quartz cuvettes with a path length of 1 cm. Reaction mixtures were prepared by pipetting the appropriate volume from stock solutions. Assays were carried out at 25 °C. All kinetic assays where carried out in a tris buffer (100 mM NaCl, 50 mM KCl, 10 mM $MgCl_2$ and 50 mM Tris pH 8.0). The assay was initiated after the solutions were equilibrated in temperature and a steady baseline was observed. Initial rates were determined by a linear least square fit of the UV trace from 20 sec to 50 sec after initiation. Rate measurements were either as single point measurements or measured as duplicates. Fitting of the rate data to models was carried out using the GraphPad Prism 6 program.

Assay conditions in the physiological/forward direction: A typical assay contained: PRPP (0-250 µM), ATP (250-750 µM), E. coli PPase (5 x the concentratrion of ATP-PRTase enzyme), ATP-PRTase enzyme (50-70 nM *L. lactis* ATP-PRTase, 160-180 nM *C. jejuni* ATP-PRTase, 280-300 nM *M. tuberculosis* ATP-PRTase), and assay buffer to make up 1 mL reaction volume. Reactions were initiated by the addition of PRPP, or sometimes ATP-PRTase enzyme after equilibration.

Assay conditions in the reverse direction typical assays contained: PP_i (0-2 mM), PRATP (0-200 µM PRATP), ATP-PRTase enzyme (50-70 nM *L. lactis* ATP-PRTase, 160-180 nM *C. jejuni* ATP-PRTase, 280-300 nM *M. tuberculosis* ATP-PRTase), and assay buffer to make up 1 mL reaction volume. Reactions were initiated by the addition of PP_i, or in some cases PRATP.

8.1.10 Determination of PRPP and PRATP concentration

Reactions were carried out as outlined in section 8.1.9. The concentration of PRPP was determined by the addition of a limiting amount of PRPP and an excess of ATP at least 20 times the concentration of that expected for PRPP. The reaction was initiated by the addition of PRPP and the reactions were allowed to go to completion (stable baseline). The change in absorbance ($\Delta_{reaction}$) was corrected for the increase in absorbance by addition of PRPP (Δ_{PRPP}) in absence of enzyme($\Delta_{absorbance} = \Delta_{reaction} - \Delta_{PRPP}$). Normal Lamber Beer law was then used to calculate the concentration in the reaction assay. The concentratrion of the stock solution was then back calculated from this. The molar extinction coefficient of 3600 M^{-1} cm^{-1} was used to calculate the concentration.[56]

The concentration of PRATP was either determined by standard kinetic assays or on a Nanodrop® ND-1000 spectrometer blanked with the appropriate buffer measuring at a wavelength of 290 nm and ensuring a pH of 8.0. Lambert Beer law was used to calculated the concentration (molar extinction coefficient of 3600 M^{-1} cm^{-1}). Determination of PRATP by standard kinetic assay conditions was carried out similar to the method described for PRPP. To ensure completion of the reaction pyrophosphate was added in access (1 mM to 1.5 mM) over the PRATP substrate. The reaction was initiated by the addition of enzyme and the reactions were allowed to go to completion (stable baseline). again the change in absorbance was corrected for the addition of the enzyme.

All concentration determinations were carried out in duplicates or triplicates.

8.1.11 Purification of *E. coli* PPase

E. coli PPase/pDEST17/E.coli Bl21* (DE3) cell cultures were grown in autoclaved LB (Lennox LB) at 20 g/L containing the appropriate antibiotics (Ampicillin, chloroamphenicol). Pre-cultures in LB (Lennox LB) media containing the appropriate antibiotics were inoculated using a single culture grown from a glycerol stock and grown overnight with shaking (180 rpm) at 37 °C. Approximately 20 mL of pre-culture per litre of cell culture were used to inoculate growths (OD_{600} of 0.1 AU). Cells were grown at 37 °C until mid-logarithmic phase (OD_{600} of 0.4–0.6 AU) and protein expression was induced by addition of isopropyl β-D-1-thiogalactopyranoside (IPTG, 0.5 mM final concentration). Growth was continued for an additional 4 hours and the cells then harvested by centrifugation (14,000 g for 15 min at 4 °C). The cell pellet was flash-frozen in liquid nitrogen and stored at −80 °C. Prior to use the cell pellet was thawed and then resuspended in low imidazole buffer (50 mM Na_2HPO_4, 500mM NaCl, 5mM $MgCl_2$ and 20mM imidazole at pH 8.5), lysed for 3x 5min 70 % power at 30% pulsation, and then spun down for 30 min at 40,000 *g*. The *E. coli* PPase enzyme was in the insoluble fraction at this stage. The supernatant was discarded and the pellet dissolved in 8 M urea, left to stand on ice for 20 min after which it was diluted 10 times. This was spun down again for 30 min at 40,000 *g*. The supernatant was then filtered and loaded onto a 5 mL GE Healthcare His trap HP column. Protein which bound to the column was eluted with high imidazole buffer (50 mM Na_2HPO_4, 500mM NaCl, 5mM $MgCl_2$ and 500mM imidazole at pH 8.5). *E. coli* PPase would elute as two peaks with the first peak containing soluble active *E. coli* PPase, the second peak precipitated out after a few minutes. Fractions containing

active enzyme were pooled together and buffer exchanged into a Tris buffer (50mM Tris, 100mM NaCl, 5mM $MgCl_2$ at pH 8.0). The enzyme was then concentrated to 4 mg/mL aliquoted out into 50 µL portions and flash-frozen in liquid N_2 and stored at -80 °C.

8.1.12 Protein structure and molecular model figures

The Protein structure and molecular model figures were all generated using PyMOL Molecular Graphics System version 1.7.4.0 from Schrodinger©, LLC. Density maps represent the unbiased ligand density ($F_O - F_C$) contoured at 3.0.

8.1.13 Polymerase chain reaction (PCR)

PCR reactions were carried out using a Veriti 96 well thermal cycle AB applied biosystems. The PCR reactions were carried out using the In-Fusion® Cloning kit with Phusion® High-Fidelity DNA Polymerase (Clonetech ®). Reactions were set up on a 50 µL scale according to Table 8.1. The amplification protocol outlined in Table 8.2 was used for the PCR reaction. Products were purified by agarose gel electrophoresis.

Table 8.1. The reaction components for PCRs.

Component	Per 50 µL reaction
Buffer 5x	10 µL
dNTP mix (10 mM of each dNTP)	1 µL
FWD Primer	2.5 µL
REV Primer	2.5 µL
Template DNA	1 µL
Phusion DNA Polymerase	0.5 µL
Nuclease free water	32.5 µL

Table 8.2. Touch down PCR protocol.

Stage	Temperature (°C)	Time (min)	Cycles
1	98.0	0.30	1 x
2	98.0	0.10	20 x
	70.0	0.30	
	72.0	0.45	
3	98.0	0.10	15 x
	45.0	0.30	
	72.0	0.45	
4	72.0	10.0	1 x

8.1.14 Colony PCR

Colony PCR reactions were carried out using a Veriti 96 well thermal cycle AB applied biosystems. The PCR reactions were carried out using taq DNA Polymerase (Invitrogen®). Reactions were set up on a 25 µL scale according to table 8.3. Template DNA was added as a

small amount of a single colony. The amplification protocol was according to manufacturer's instructions.

Table 8.3. The reaction components for PCRs.

Component	Per 50 µL reaction
Standard taq buffer	2.5 µL
dNTP mix (10 mM of each dNTP)	0.5 µL
T7 FWD Primer (10 µM)	0.5 µL
T7 REV Primer (10 µM)	0.5 µL
Template DNA	1 colony
Taq DNA Polymerase	0.125 µL
Nuclease free water	20.88 µL

8.1.15 Agarose gel electrophoresis

Agarose gel electrophoresis was carried out on a Mini-Sub® Cell GT (Bio-Rad). Visualisation and photographs of DNA were done using a Molecular Imager® Gel Doc™ XR (Bio-Rad). Agerose gel was prepared by mixing 0.3 gram of agarose with 29.4 mL of Milli-Q water and 600 µL 50 X TAE buffer (2.5 M Tris, 50 mM EDTA, 1 M acetic acid). Agarose was dissolved by heating using a microwave until fully dissolved. The gel was allowed to cool to approximately 60 °C in the oven (approximately 20 min) followed by the addition of 3 µL SYBSafe dye. After mixing the gel was poured into a sealed gel tray including a comb and left to set. 1 portion loading dye (60 mM Tris-HCl, 60 mM EDTA, 0.2% (w/v) orange G, 0.05% (w/v) xylene cyanol, 60% (v/v) glycerol) was mixed with 5 portions sample and loaded on the gel. A DNA ladder was also included for size identification purposes. The electrophoresis was performed for 45 min at 80 V.

8.1.16 Transformation

Appropriate chemically competent cells (50 µL) were carefully thawed on ice followed by the addition of purified plasmid (75 – 120 ng/µL). The cells were left to stand on ice for 30 min then placed on a heat block at 42 °C for 30-45 seconds and then back on ice for 2 minutes. To this 250 µL of SOC medium (2% (w/v) tryptone, 0.5% (w/v) yeast extract, 10 mM NaCl, 2.5 mM KCl, 10 mM MgSO4, 20 mM glucose. Filter sterilised) was added and left to grow at 37 °C for 1 hour. The transformed cells (20 µL to 100 µL) were plated out on LB-agar containing the appropriate antibiotic(s) and left to grow overnight at 37 °C.

8.1.17 Plasmid extraction and purification

Cells used for plasmid extraction and purification were grown overnight as 5 mL cultures in presence of the appropriate antibiotic(s). Typically 1 mL of culture was processed according to procedures provided with the extraction kit. (ZR Plasmid miniprep^{TM} –classic, Zymo Research) concentration of the purified plasmid was established by measuring the absorbance at 260 nm an on appropriately blanked Nanodrop ND-1000 spectrophotometer.

8.1.18 Sequencing

The verification of the DNA sequence was performed after subcloning into the destination vector. Candidates for sequencing were identified by colony PCR as described in Section 8.1.14. Sequencing was carried out by Massy university, Palmerston north NZ. Samples were prepared with 60-125 ng and 4 pM T7 FWD or REV primers made up to 20 µL with nuclease free water.

8.1.19 Glycerol stocks

Glycerol stocks were made from cell cultures grown overnight at 37 °C in presence of the appropriate antibiotics. 800 µL of cell culture was mixed with 400 µL of autoclaved glycerol (50% v/v) stock solution and flash frozen by submerging in liquid nitrogen. Stocks were made either in a 1.6 mL sterile Eppendorf or screw-top tube.

8.1.20 General procedure protein expression in autoinduction media

The method developed by Studier was used for autoinduction of protein expression.[198] general description for a 1 liter culture: To 960 mL of water was added 10 g tryptone and 5 g yeast extract and autoclaved in a 3 liter baffled flask. To this 20 mL 50 X M stock solution (1.25 M Na2HPO4, 1.25 M KH2PO4, 2.5 M NH4Cl, 0.25 M Na2SO4. Autoclaved), 20 mL 50 X 5052 stock solution (25% (w/v) glycerol, 2.5% (w/v) glucose, and 10% (w/v) lactose, filter sterilised),0.2 mL trace metal solution (50 mM $FeCl_3$, 20 mM $CaCl_2$, 10 mM $MnCl_2$/$ZnSO_4$, 2 mM $CoCl_2$, 2 mM $CuCl_2$, 2 mM $NiCl_2$, 2 mM $NaMoO_4$, 2mM Na_2SeO_3, and 2mM H_3Bo_3 in 60 mM HCl) and 1 mL of 1M $MgSO_4$ was added. This culture medium was inoculated with pre-culture to yield a starting OD_{600} of 0.1 AU. Cultures were grown at 37 °C for 20 hours with shaking (180 rpm).

8.1.21 antibiotics

Antibiotics used were added from a 1000 times concentrated stock. Stocks were prepared in batch and stored at -80 °C. Sufficient stock was added to achieve final concentrations of 100 µg/mL ampicillin, 50 µg/mL kanamycin, 50 µg/mL spectinomycin and chloramphenicol 25 µg/mL.

8.1.22 Cell lysis

Cell lysis was carried by sonication on a Omni-Ruptor 4000 Ultrasonic Homogenizer from Omni International. Typically the cell pellet was resuspended in ice cooled lysis buffer (20 mL to 100 mL) and 2 µL Benzonase nuclease® (Novagen) was added to break down DNA. Sonication was performed in a glass beaker surrounded by packet ice generally for 3 cycles of 5 min at 30 % to 40 % pulsation and 70 % power.

The soluble fraction was separated by centrifugation at 40,000 g for 30 min at 4°C, transferred to a clean tube and spun for an additional 15 min.

8.2 Experimental procedures for chapter 2

8.2.1 Cloning of M. tuberculosis ATP-PRTase

The cloning of *M. tuberculosis* ATP-PRT made use of the In-Fusion® HD Cloning Kit (Clontech® Laboratories Inc.). Primer design was carried out by Dr G. Mittlestädt according to published procedures. In the FWD primer the thrombin cleavage site was replaced by a TEV recognition site. PCR was carried out according to Section 8.1.13, PCR fragments were assessed for appropriate size by agarose gel electrophoresis and ligated into vector pET28a.

MtuHisGin28aFWD: CGCGCGGCAGCCATATGGAAAACCTGTATTTTCAGGGCAGCGGCGCGATG CTGCGGGTCGCGGTTCC

MtuHisGin28aREV: GGTGGTGGTGCTCGAGTCATCAGAATCGGCAGAACCTGAT

8.2.2 Purification of *Campylobacter jejuni* ATP-PRTase

Expression and purification of *C. jejuni* ATP-PRTase was carried out according to established procedures by Dr G. H. mittlestäd. *C. jejuni* ATP-PRTase/pDEST17/*E.coli* Bl21* (DE3) pBB540/pBB542 (Chaperone 3) cell cultures were grown in autoclaved LB (Lennox LB) at 20 g/L containing the appropriate antibiotics (Ampicillin, chloroamphenicol and spectinomycin). Pre-cultures in LB (Lennox LB) media containing the appropriate antibiotics

were inoculated using a single culture grown from a glycerol stock or on an agar plate and grown overnight with shaking (180 rpm) at 37 °C. Approximately 20 mL of pre-culture per litre of cell culture were used to inoculate growths (OD_{600} of 0.1 AU). Cells were grown at 37 °C until mid-logarithmic phase (OD_{600} of 0.4–0.6 AU). The temperature was lowered to 23 °C before induction of protein expression by addition of isopropyl β-D-1-thiogalactopyranoside (IPTG, 0.5 mM final concentration), and growth was continued overnight. Cells were harvested by centrifugation (14,000 *g* for 15 min at 4 °C) and the cell pellet was flash-frozen in liquid nitrogen and stored at −80 °C for purification at a later stage or used the same day. The cell pellet was resuspended in low imidazole buffer (50 mM Na_2HPO_4, 500 mM NaCl, 5 mM $MgCl_2$ and 20 mM imidazole at pH 8.5), lysed for 3x 5min 80 %, spun down for 30 min at 40,000 *g*. The supernatant was filtered and loaded onto a 5 mL GE Healthcare His trap HP column. Protein which bound to the column was eluted with high imidazole buffer (50 mM Na_2HPO_4, 500 mM NaCl, 5 mM $MgCl_2$ and 500 mM imidazole at pH 8.5) and fractions containing enzyme were pooled together. Pooled fractions were buffer exchanged into a Tris buffer (50 mM Tris, 100 mM NaCl, 5 mM $MgCl_2$ at pH 8.5) and treated with TEV protease at a 1% mol/mol concentration at 4°C for three days. TEV treated enzyme was run through a 26/60 size exclusion column into buffer containing 50 mM Tris and 100mM NaCl at pH 8.5. Fractions containing *C. jejuni* ATP-PRTase were concentrated and 50 µL aliquots at a concentration of 10-12 mg/mL, flash-frozen in liquid N_2 and stored at -80 °C. 2 litres of culture would yield approximately 20-30 mg of protein per purification

8.2.3 Purification of ATP-PRTase from *Mycobacterium tuberculosis*

Expression and purification of *M. tuberculosis* ATP-PRTase was carried out according to published procedures.[47,70] *M. tuberculosis* ATP-PRTase/pET28a/ *E. coli* BL21 (DE3) Star pLysS cell cultures were grown in autoinduction media (Methods 8.1.20)containing Kanamycin. Pre-cultures (50-100 mL) in LB (Lennox LB) media containing Kanamycin were inoculated using a single culture as a glycerol stock or from an agar plate and grown overnight with shaking (180 rpm) at 37 °C. Approximately 40 mL of pre-culture per litre of cell culture were used to inoculate growths (OD_{600} of 0.1 AU). Cells were grown with shaking (180 rpm) at 37 °C for 1 day. Cells were harvested by centrifugation (14,000 g for 15 min at 4 °C). The cell pellet was resuspended in low imidazole buffer (50 mM Na_2HPO_4, 500mM NaCl and 10mM imidazole at pH 7.8), lysed for 3x 5min 80% power and 30% pulse, spun down for 30 min at

40,000 *g*, supernatant transferred to a second tube and spun for a further 15 min at 40,000 *g*. The supernatant was filtered and loaded onto a 5 mL GE Healthcare Hitrap TALON® crude column. Protein which bound to the column was eluted with high imidazole buffer (50 mM Na_2HPO_4, 500 mM NaCl, and 250 mM imidazole at pH 7.8) and fractions containing enzyme were pooled together. Pooled fractions were buffer exchanged into a Tris buffer (50 mM Tris, 100 mM NaCl, 5 mM $MgCl_2$ at pH 7.8) and treated with TEV protease at a 1% mol/mol concentration at 37 °C for 2-3 hours. TEV treated enzyme was run through a 26/60 size exclusion column into Tris buffer. Fractions containing *M. tuberculosis* ATP-PRTase were concentrated to 15-20 mg/mL and divided into 50-100 µL aliquots and either used within the next week or flash-frozen in liquid N_2 and stored at -80 °C. A 1 litre culture typically yielded approximately 100 mg of pure active protein.

8.2.4 Crystallography

8.2.4.1 Data collection and processing

Single crystal X-ray diffraction data was collected at the Australian Synchrotron on the MX1 and MX2 beamlines running the Blu-Ice software. XDS, iMosflm, and aimless were used for data processing. [199–201] Resolution cut-off was based on CC1/2 > 0.5 values.[202]

8.2.4.2 Crystal conditions for *M. tuberculosis* ATP-PRTase

Crystals were grown according to published conditions for the growth of APO crystals (1.8 M $MgSO_4$, 0.1 M 2-(N-morpholino)ethanesulfonic acid (MES) pH 6.5) with the addition of 10 mM ATP. Crystals were obtained by mixing equal volume of 15-20 mg/mL protein with 90-100% condition (1-2µL each) using the hanging drop vapour diffusion method. Crystals formed between 12 to 48 hours. Crystals grew as uniform triangular prisms of various size. Protein crystals were cryoprotected by dipping the crystal in a second droplet of condition including 20% glycerol.

8.2.4.3 Structure solution *M. tuberculosis* ATP-PRTase

Initial phases were solved by a molecular replacement strategy using the CCP4 program suite.[201] The initial search model was generated from the APO *M. tuberculosis* ATP-PRT (PDB code: 1NH7). The ligands and water molecules were removed from the search model. Refinement of the structure was done using multiple rounds of refinement using coot.[203] The structure was validated using the MolProbity server and wwPDB validation server.[204,205]

8.2.4.4 Crystal conditions for *C. jejuni* ATP-PRTase

Initial crystallisation screens were performed with mosquito® Crystal robot (TTP Labtech), and the JCSG-plus™ HT-96, Clear Strategy™ I HT-96 and Clear Strategy™ II HT-96 crystallisation screens (Molecular Dimensions), and 96-well sitting drop iQ plates (TTP Labtech). Each compartment contained 40 µL condition which was mixed 1/1 v/v with protein to give a final drop volume of 800 nL. Screens were set up with freshly purified protein premixed with either ATP or His/AMP. A final protein concentration of 3.5 mg/mL in the drop was used. This yielded a number of promising conditions which were further optimised. Optimisations were carried out on a 24-well VDX plates (Hampton Research) using hanging-drop vapour diffusion at 20 °C. Wells would contain 500 µL condition and protein was mixed 1/1 with condition. The His/AMP crystals were put down at 2 µL and the ATP crystals were put down at 4 µL. Slow addition of an equil volume of 40 % glycerol in conditions to the crystal drop prior to crystal looping provided cryoprotection

Screening condition	Ligand concentration	Refined condition
Clear strategy™ I HT-96 E9	1mM Histidine 5mM AMP	0.1 M Tris pH 7.5 0.2 M $MgCl_2$ 13-15 % w/vPEG 4000
Clear strategy™ I HT-96 A9	10 M ATP	0.1 M AcONa pH 5.5 0.1 M $MgCl_2$ 13-15 % w/vPEG 4000

8.2.4.5 Structure solution *C. jejuni* ATP-PRTase

Initial phases were solved by a molecular replacement strategy using the CCP4 program suite.[201] The initial search model was generated from *C. jejuni* ATP-PRTase with histidine bound (PDB code: 4YB5). For the His/AMP bound structure the search model contained the full hexamer which gave a single solution, further refinement of the structure was done using coot.[203] The search model for the ATP bound structure containing only domain I and II. A suitable dimeric unit was identified and used in subsequent runs until two hexameric arrangement were found. The missing regulatory domains were added through multiple

rounds of refinement using Coot. Both structures were validated using the MolProbity server and wwPDB validation server before deposition.[204,205]

8.3 Experimental procedures for chapter 3

8.3.1. Expression and purification of *M. tuberculosis* PRPPase.

Plasmids encoding the *M. tuberculosis* PRPPase gene were kindly provided by Prof V. L. Schramm. *M. tuberculosis* PRPPase/E.coli Bl21 (DE3) pBB540/pBB542 (Chaperone 3) cell cultures were grown in autoclaved LB (Lennox LB) at 20 g L^{-1} containing the appropriate antibiotics (Kanamycin , chloramphenicol and spectinomycin). Pre-cultures in LB (Lennox LB) media containing the appropriate antibiotics were inoculated using a single culture grown from a glycerol stock or on an agar plate and grown overnight with shaking (180 rpm) at 37 ^{o}C. approximately 45 mL of pre-culture per litre of cell culture were used to inoculate growths (OD_{600} of 0.1 AU). Cells were grown at 37 ^{o}C until mid-logarithmic phase (OD_{600} of 0.4–0.6 AU). The temperature was lowered to 23 ^{o}C before induction of protein expression by addition of isopropyl β-D-1-thiogalactopyranoside (IPTG, 0.1 mM final concentration), and growth was continued overnight. Cells were harvested by centrifugation (14,000 g for 15 min at 4 ºC) and the cell pellet was either used the same day or flash-frozen in liquid nitrogen and stored at −80 ^{o}C for purification at a later date. The cell pellet of 2 liters of cell culture was resuspended in 40mL of low imidazole buffer (50 mM Na_2HPO_4, 300 mM NaCl and 20mM imidazole at pH 8.5), lysed for 3x 5min 80%, spun down for 30 min at 40,000 *g*. The supernatant was filtered and loaded onto a 5 mL GE Healthcare His trap HP colomn. Protein which bound to the column was eluted over a 4 CV gradient with high imidazole buffer (50 mM Na_2HPO_4, 500mM NaCl, 5mM $MgCl_2$ and 450mM imidazole at pH 8.5) resulting in two elution peaks. The second peak corresponded to the protein of interest and fractions containing enzyme were pooled together (Figure 8.1). Pooled fractions were buffer exchanged into a tris buffer (50 mM Tris, 100mM NaCl, 5 mM $MgCl_2$ at pH 8.5). The protein was then concentrated to approximately 5 mg/mL and devided into 50 µL aliquots, flash-frozen in liquid N_2 and stored at -80 °C. 2 litres of culture would yield approximately 50 mg of protein per purification.

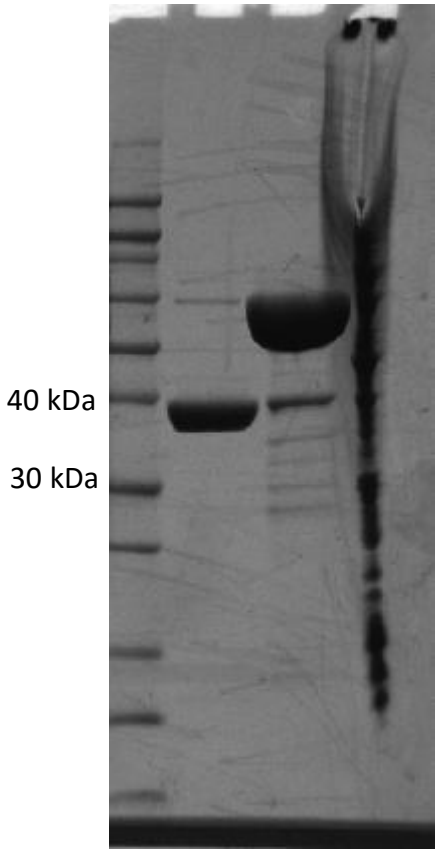


Figure 8.1. PRPPase SDS page gel. Lane 1 cell pellet, lane 2 Peak 1, Lane 3 the target protein as peak 2.

8.3.2 Purification of ribose kinase

Plasmids encoding the His tagged gene were kindly provided by Prof V. L. Schramm. Ribose kinase/E.coli TOP10 cell cultures were grown in autoclaved LB (Lennox LB) at 20 g L^{-1} containing 50 µg/mL ampicillin. Pre-cultures in LB (Lennox LB) media containing the appropriate antibiotics were inoculated using a single culture grown from a glycerol stock and grown overnight with shaking (180 rpm) at 37 ^{o}C. approximately 20 mL of pre-culture per litre of cell culture were used to inoculate 1L growths (OD_{600} of 0.1 AU). Cells were grown at 37 ^{o}C until mid-logarithmic phase (OD_{600} of 0.4–0.6 AU). Protein expression was induced with 0.2 % L(+)-arabinose (2 g/L) final concentration, and growth was continued for at least 4 hours. Cells were harvested by centrifugation (14,000 g for 15 min at 4 °C) and the cell pellet washed with cold buffer (50 mM Tris HCl, 300 mM NaCl, at pH 7.8), centrifuged another time (14,000 g for 15 min at 4 ºC). The cell pellet of 1 liter of cell culture was resuspended in 20mL of Tris buffer (50 mM Tris HCl, 300 mM NaCl at pH 7.8), lysed for 3x 5min 80%, spun down for 30 min at 40,000 *g*. The supernatant was filtered and loaded onto

a 5 mL GE Healthcare His trap HP colomn. Protein which bound to the column was eluted over an 7CV gradient with RK eluting as a single peak with high imidazole buffer (50 mM Tris HCl, 300 mM NaCl, and 450 mM imidazole at pH 8.5). Fractions containing enzyme were pooled together (Figure 8.2). Pooled fractions were buffer exchanged into a low imidazole buffer (50mM Tris, 100mM NaCl, 5mM $MgCl_2$ at pH 8.5). The protein was then concentrated to approximately 10 mg/mL and devided into 50 µL aliquots, flash-frozen in liquid N_2 and stored at -80 °C. 1 litres of culture would yield approximately 70 mg of protein per purification.

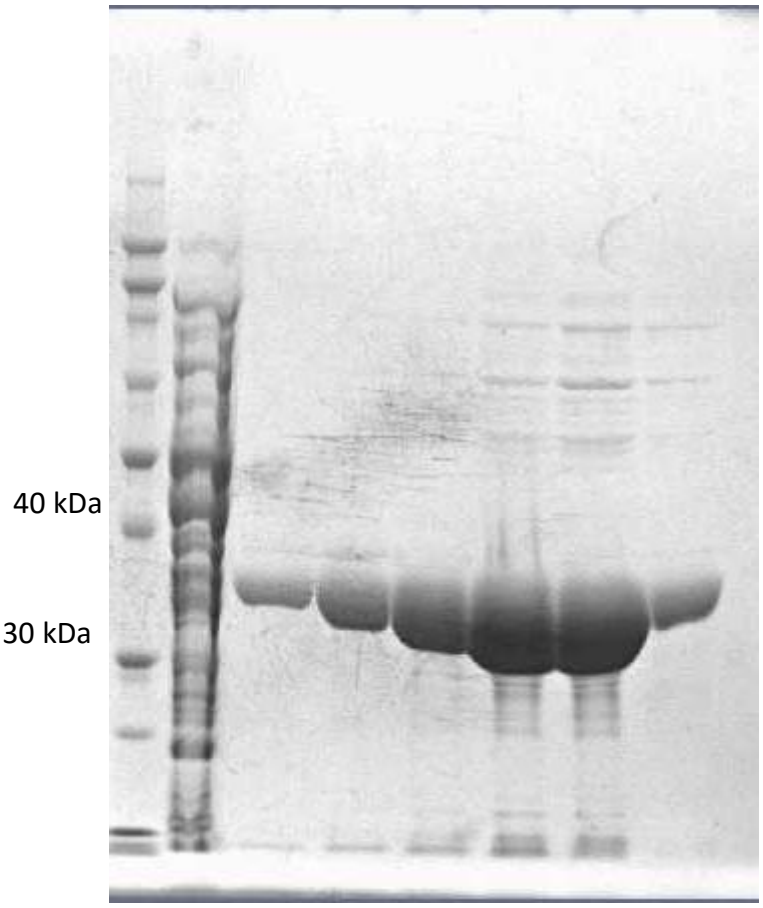
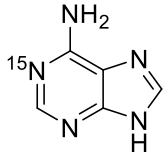


Figure 8.2. Gel image of fraction containing ribokinase.

8.3.3 Methods

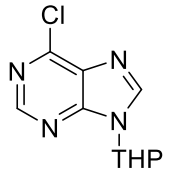
8.3.3.1 synthesis of [1-^{15}N]adenine:



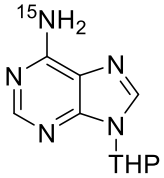
[1-^{15}N]adenine 3.6: A mass spectrometry vial containing a stir bar and 4-amino-imidazole-5-carbonitrile (0.49 mmol, 53 mg) was flushed with a gentle

flow of nitrogen gas for 30 min. 0.5 mL of ^{15}N formamide (Cambridge Isotope Laboratories) was added and the 4-amino-imidazole-5-carbonitrile was fully dissolved before heating on a preheated oil bath at 180 °C for 2.5 hours with active stirring. The reaction was allowed to cool down upon which solids formed. The crude reaction mixture was further cooled on ice for 30 min and solids filtered over a cotton wool plug. The excess ^{15}N formamide was recovered. The solids were washed 2 X with ice cold water to give the final [1-^{15}N]adenine product as a yellow powder which was used without further purification. (35 mg, 52% yield). The spectroscopic analysis was in agreement with reported literature for this structure.[100] IR (v): 537.6, 633.1, 720.3, 793.6, 843.1, 903.5, 936.3, 1123.3, 1246.2, 1305.0, 1365.7, 1416.3, 1448.9, 1591.1, 1668.5, 2788.4, 3107.4 cm^{-1}): ^{1}H NMR (400 MHz, DMSO-D6); δ 7.06 (bs, 2H, NH_2), 8.07 (s, 1H, H8), 8.09 (d, *J* = 15.3 1H, H2); HRMS(ESI) *m/z* calcld. for $[C_5H_6N_4{}^{15}N]^+$: 137.0588 obsd.: 137.0583.

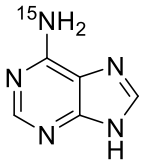
8.3.3.2 [6-^{15}N]Adenine synthesis



6-Chloro-N9-(2-tetrahydropyranyl)-purine (3.9): 6-chloro-purine (1.94 mmol, 300mg) was heated to 60°C in 7 mL of ethyl acetate. *Pare*-toluenesulphonic acid (0.1 mmol, 20 mg) was added and the suspension left to stir for an additional 20 min. 2,3-dihydro-2*H*-pyran (2.52 mmol, 230 µL) was dropwise added over 5-10 min. after approximately 10 min all solids were dissolved and the resulting mixture left to stir for an additional 1.5 hours. t.l.c. analysis (CH_2Cl_2/MeOH, 95/5, v/v) showed formation of a higher running product (Rf 0.4). Concentrated ammonia was added, 0.2 mL, left to stir for an additional 5 min and the organic phase washed 2 X with water and 1x with brine, dried over anhydrous $MgSO_4$ and filtered. Purification by flash chromatography (silica gel, CH_2Cl_2/MeOH, 96/4, v/v) gave the title compound 3.5 as a white solid at 93% yield (430 mg, 1.8 mmol). The product was in good agreement with the reported spectra.[103] IR (v): cm^{-1}): ^{1}H NMR (400 MHz, DMSO-D6); δ 1.56-1.62 (m, 2H, 2xH4'), 1.68-1.78 (m, 1H, H3'*b*), 1.79-2.02 (m, 2H, H2'*b*, H3'*a*), 2.27-2.37 (m, 1H, H2'*a*), 3.67-3.74 (m, 1H, H5'*b*), 3.99-4.02 (m, 1H, H5'*a*), 5.77 (dd, *J* = 2 Hz, *J* = 10.8 Hz, 1H, H1'), 8.78, 8.89 (2x s, 2x 1H, H2, H8); ^{13}C NMR (101 MHz, DMSO-D6); δ 22.64 (C2'), 24.87 (C4'), 30.12 (C3'), 68.20 (C5'), 82.14 (C1'), 131.40, 146.00, 149.74, 151.74, 152.24 (C2, C4, C5, C6, C8).



N9-(2-Tetrahydropyranyl)-[6-15N]-adenine (3.10): To a Kevlar® coated metal pressure reactor 6-Chloro-N9-(2-tetrahydropyranyl)-purine 3.5 (0.1.27 mmol, 300 mg), $NaHCO_3$ (3.77 mmol, 318 mg), and $^{15}NH_4Cl$ (2.5 mmol, 139 mg) was added. DMSO (1.9 mL) was added and the reactor sealed immediately after this to enclose the released ammonia. The reactor was heated at 80 °C for 4.5 days. The resulting suspension was filtered through celite and concentrated on a rotary evaporator (1 day). Purification by flash chromatography (silica gel, CH_2Cl_2/MeOH, 99/1 → 96/4, v/v) gave the title compound 3.3 as a white solid at 84% yield (430 mg, 1.8 mmol). The product was in good agreement with the reported spectra.[103] IR (v): 449.8, 485.0, 544.4, 567.6, 634.7, 648.7, 675.3, 781.8, 794.9, 821.6, 856.2, 877.2, 904.5, 948.1, 1040.5, 1082.6, 1108.9, 1143.5, 1179.1, 1207.0, 1221.9, 1266.1, 1335.7, 1396.3, 1442.5, 1483.8, 1563.3, 1592.7 cm^{-1}: ^1H NMR (400 MHz, DMSO-D6); δ 1.60-1.80 (m, 3H, H3'*b*, 2x H4'), 1.96-2.11 (m, 3H, 2xH2', H3'*a*), 3.75 (t, *J* = 11.7 Hz, 1H, H5'*b*), 4.14 (d, *J* = 12.0 Hz, 1H, H5'*a*), 5.69 (d, *J* = 10.8 Hz, 1H, H1'), 6.31 (d, *J* = 90.0 Hz, 2H, $^{15}NH_2$), 8.01, 8.34 (2x s, 1x1H, H2, H8); ^{13}C NMR (101 MHz, DMSO-D6); δ 22.79 (C3'), 24.87 (C4'), 31.87 (C2'), 68.77 (C5'), 81.85 (C1'), 138.23, 149.24, 152.85, 155.52, 155.72 (C2, C4, C5, C6, C8).



[6-^{15}N]-Adenine (3.8). N9-(2-Tetrahydropyranyl)-[6-15N]-adenine (3.6) (1.07 mmol, 234 mg) was dissolved in dioxane (3.8 mL) and 7.6 mL 0.01 M HCl in water added. The pH was further adjusted to 2.0 by the addition of 20 drops of concentrated HCl. After addition of acid a cloudy solution is formed which went clear after 15 min of additional stirring. The reaction was left overnight and then the volume reduced to 20 mL. an additional 20 mL of water was added and the volume reduced to 20 mL. This step was repeated 3 times and the solution further neutralised by the addition of $NaHCO_3$ (~pH 7.0 by pH paper) and concentrated. This gave [6-^{15}N]adenine as the only observed product. The adenine product was in good agreement with the reported spectra.[103] IR (v): 537.6, 632.9, 720.2, 793.5, 843.3, 903.4, 936.1, 1123.4, 1246.1, 1304.8, 1365.7, 1416.3, 1589.8, 1668.2, 2785.7, 3076.6 cm^{-1}: ^1H NMR (400 MHz, DMSO-D6); δ 7.06 (d, *J* = 89.6 Hz, $^{15}NH_2$), 8.07, 8.09 (2x s, 2x 1H, H2, H8); ^{13}C NMR (101 MHz, DMSO-D6); δ 117.56, 140.44, 152.62, 155.36, 155.57 (C2, C4, C5, C6, C8). HRMS(ESI) *m/z* calcld. for $[C_5H_6N_4{}^{15}N]^+$: 137.0588, obsd.: 137.0594.

8.3.4 Coupled enzymatic synthesis of PRATP from ribose

Reactions were set up according to the condition outline in table 8.4, generally on a 12.5 mL scale. All reaction components were combined except for *C. jejuni* ATP-PRTase and *E. coli* PPase. The reaction was left at 37 °C for 30 min after which the remaining two enzymes were also added. The reaction was allowed to stand for an additional 90 min before acidification by the addition of 1M HCl until pH 6.0 by litmus paper. The solution was frozen in liquid nitrogen and lyophilised. The resulting white powder was taken up in a minimal amount of buffer (50 mM $AcONH_4$ pH 6.0), spun down for 10 min at 14,000 *g* to remove solids, and filtered through a 0.45 micron syringe filter. The reaction mixture was loaded onto a P2 (Bio-Rad®)size exclusion gel made up in a XK 16/40 column. The PRATP product eluted together with the rearranged 6-*N*-PRATP product. Fraction containing PRATP were combined, frozen in liquid nitrogen and lyophilised overnight. The resulting product was used without further purification.

Table 8.4. Components of the couled enzymatic synthesis of PRATP from ribose.

Component	Final concentratrion
Na_2HPO_4	50 mM
NaCl	100 mM
$MgCl_2$	10 mM
PEP (Na salt)	20 mM
D -ribose	2.0 mM
ATP	2.2 mM
Myokinase	5 U/mL
Pyruvate Kinase	20 U/mL
Ribokinase	40 µg/mL
PRPPase	32 µg/mL
C. jejuni ATP-PRTase	1.1 µM
PPase	1.1 µM

8.3.5 Coupled enzymatic synthesis of isotopically labelled PRATP from ribose

Reactions were set up according to the condition outline in table 8.5, generally reactions were carried out on a 500 µL scale. All reaction components were combined except for *C. jejuni* ATP-PRTase and *E. coli* PPase. The reaction was left at 37 °C for 30 min after which the remaining two enzymes were added. If the reaction involved the ^3H isotopic label additional unlabelled D-ribose (0.5 mM) was added prior to the addition of the *C. jejuni* ATP-PRTase and *E. coli* PPase and incubated for an additional 15 min before the remaining two enzymes were also added. The reaction was then monitored by HPLC according to the general

method described in Section 8.3.3. The reaction was quenched by the addition of 100 µL 0.1M H_2SO_4, filtered through a spin concentrater at 4 °C and purified according to the general procedure described in section 8.3.7.

Table 8.5. Components of the coupled enzymatic synthesis of isotopically labelled PRATP from ribose.

Component	Final concentratrion
Isotopically labelled D-ribose	20 µCi to 50 µCi
Na_2HPO_4	50 mM
NaCl	100 mM
$MgCl_2$	10 mM
PEP (Na salt)	20 mM
ATP	0.7 mM
Myokinase	5 U/mL
Pyruvate Kinase	20 U/mL
Ribokinase	40 µg/mL
PRPPase	32 µg/mL
C. jejuni ATP-PRTase	1.1 µM
PPase	1.1 µM

8.3.6 Preparation of HPLC eluents

Aqueous HPLC buffers were prepared with Milli-Q water and adjusted in pH were made as per Section 8.1.5. The aqueous buffer component was filter sterilised before the organic solvent component was added. Buffers were used within a week and stored at 4 °C when not in use. Phosphate buffers were made by mixing the appropriate mono- and di-basic phosphate salts as a 100 mM stock solution until the desired pH was achieved.

- Ion-paring buffer: Buffer A: 50 mM KH_2PO_4, 8 mM (*n*-Bu)$_4$$NHSO_4$ at pH 5.0, Buffer B; 50% v/v buffer A/Acetonitrile
- Desalting buffer PRATP: Buffer A: 50 mM $AcONH_4$ at pH 5.8, Buffer B: 50% v/v buffer A/Acetonitrile
- Desalting buffer ATP: Buffer A: 0.1% formic acid, Buffer B: 50% v/v buffer A/Acetonitrile

All organic solvents used in were of HPLC grade and used without further filtering.

8.3.7 HPLC methods

Purification of compounds were performed on a Phenomenex Kinetex C18 5u 100 Å 250-4.70 mm reverse phase column, linked to a Water Alliance 2695 HPLC with a waters 2487

Dual Absorbance Detector. Columns were equilibrated for at least 1 hour before use at a flowrate of 1 mL/min. All purifications were carried out at the same flowrate of 1 mL/min. The purification of PRATP and ATP were carried out in a two-step process by separation from the reaction mixture by the ion paring method (Section 8.3.6), PRATP eluted at 21 min and ATP at 22 min, followed by desalting (Section 8.3.6) with the appropriate eluent system, PRATP eluted at 15 min ATP eluted at 16 min.

Ion paring method:

Time (min)	Buffer A	Buffer B
0-2	100 %	0 %
2-25	0 %	100 %
25-27	0 %	100%
27-28	100 %	0 %
28 -35	100 %	0 %

Desalting method

Time (min)	Buffer A	Buffer B
0-3	100 %	0 %
3-10	0 %	100 %
10-23	0 %	100%
23-24	100 %	0 %
24 -30	100 %	0 %

8.3.8 Measuring KIEs

8.3.8.1 Condition for the measurement of KIES for *M. tuberculosis* ATP-PRTase

Kinetic isotope effects were measured under competitive conditions. An experiment would include 4 reactions of partial conversion (20%-40%), 2 100% conversion reactions and 1 negative control (Table 8.7). A master mix (Table 8.6) was made containing, equal amounts of ^{3}H and ^{14}C isotopically labelled PRATP. Per reaction enough label was added to give 45,000 CPM/reaction/per isotope so a total of 315,000 CPM/per isotope in total in the master mix (Table 8.6). Additional unlabelled PRATP was added to give a final concentration of 9-10 µM (achieving this concentration after enzyme and buffer is added to the individual reactions) after enzyme was also added. this was made up in a tris buffer (100 mM NaCl, 50 mM KCl, 10 mM $MgCl_2$ and 50 mM Tris HCl pH 8.5) and phosphonoacetic acid was added to a final concentration of 10mM.

Table 8.6. Components of the master mix for M. tuberculosis ATP-PRTase.

Descriptor	Amount/concentration
Master mix total volume	~ 1350 µL
^{3}H labelled PRATP	315,000 CPM
^{14}C labelled PRATP	315,000 CPM
PRATP	Final concentration of 9 µM
Phosphonoacetic acid	10 mM

Table 8.7. Set up of the individual reactions.

Descriptor	20-40% conversion	100% conversion	Blank
Master mix	190 µL	190 µL	190 µL
Enzyme	244 nM	16 µM	0
Buffer	8.25-8.5 µL	0 µL	10 µL
Number of RXN	4X	2X	1X

After enzyme was added the reaction was incubated at 25 °C for 17 min, 20 µL 0.1M H_2SO_4 was then added to each tube to quench the reaction. Processing of the sample was done according to the general procedure described in Section 8.3.8.4 immediately after.

8.3.8.2 Condition for the measurement of KIEs for *C. jejuni* ATP-PRTase

Kinetic isotope effects were measured under competitive conditions. An experiment would include 4 reactions of partial conversion (20%-40%), 2 100% conversion reactions and 1 negative control (Table 8.9). A master mix (Table 8.8) was made containing, equal amounts of ^{3}H and ^{14}C isotopically labelled PRATP. Per reaction 315,000 CPM/per isotope was added, additional unlabelled PRATP would be added to give a concentration of 18-19 µM in the reaction (so enough PRATP that after the addition of the enzyme and buffer the PRATP concentration of isotopically labelled and unlabelled PRATP would be 18-19 µM). This was made up in a Tris buffer (100 mM NaCl, 50 mM KCl, 10 mM $MgCl_2$ and 50 mM Tris HCl pH 8.5) and phosphonoacetic acid was added to a final concentration of 10mM in the reaction.

Table 8.8. Components of the master mix for *C. jejuni* ATP-PRTase

Descriptor	Amount/concentration
Total volume	~ 1350 µL
^{3}H labelled PRATP	315,000 CPM
^{14}C labelled PRATP	315,000 CPM
PRATP	Final concentration of 18.5 µM
Phosphonoacetic acid	10 mM

Table 8.9. Set up of the individual reactions.

Descriptor	20-40% conversion	100% conversion	Blank
Master mix	190 µL	190 µL	190 µL
Enzyme	1.7 µM	18 µM	0
Buffer	6-7 µL	0 µL	10 µL
Number of RXN	4X	2X	1X

After enzyme was added the reaction was incubated at 25 °C for 20 min, 20 µL 0.1M H_2SO_4 was then added to quench the reaction. Processing of the sample was done according to the general procedure described in Section 8.3.8.4 immediately after.

8.3.8.3 Condition for the measurement of KIEs for *L. lactis* ATP-PRTase

Kinetic isotope effects were measured under competitive conditions. An experiment would include 4 reactions of partial conversion (20%-40%), 2 100% conversion reactions and 1 negative control (Table 8.11). A master mix (Table 8.10) was made containing, equal amounts of ^{3}H and ^{14}C isotopically labelled PRATP. Per reaction 315,000 CPM/per isotope was added, additional unlabelled PRATP would be added to give a concentration of 18-19 µM in the reaction (so enough PRATP that after the addition of the enzyme and buffer the PRATP concentration of isotopically labelled and unlabelled PRATP would be 18-19 µM). This was made up in a Tris buffer (100 mM NaCl, 50 mM KCl, 10 mM $MgCl_2$ and 50 mM Tris HCl pH 8.5) and phosphonoacetic acid was added to a final concentration of 10mM in the reaction.

Table 8.10. Components of the master mix for *C. jejuni* ATP-PRTase

Descriptor	Amount/concentration
Total volume	~ 1350 µL
^{3}H labelled PRATP	315,000 CPM
^{14}C labelled PRATP	315,000 CPM
PRATP	Final concentration of 35 µM
Phosphonoacetic acid	10 mM

Table 8.11. Set up of the individual reactions.

Descriptor	20-40% conversion	100% conversion	Blank
Master mix	190 µL	190 µL	190 µL
Enzyme	150 nM	940 nM	0
Buffer	6-7 µL	0 µL	10 µL
Number of RXN	4X	2X	1X

After enzyme was added the reaction was incubated at 25 °C for 20 min, 20 µL 0.1M H_2SO_4 was then added to quench the reaction. Processing of the sample was done according to the general procedure described in Section 8.3.8.4 immediately after.

8.3.8.4 General procedure for separation of substrate and products by solid phase extraction:

Reactions were loaded onto a pre-equilibrated (2x 4mL H_2O and 2x 4mL elution buffer, 50mM $(NH_4)_2CO_3$ at pH 6.0 with AcOH) Grace Extract-Clean SPE carbo 150mg/4mL column. Columns were eluted with 4x 1 mL of elution buffer and the flowthrough collected directly into a scintallation vial. Samples were dried under speed vacuum evaporation for at least 12 hours and the resulting solids were taken up in MilliQ water (500 µL). 10 mL Ultima Gold™ Cocktail was added, here it was found important to make sure all solids had completely disolved in water before scintillation counting liquid was added. Samples were mixed for 2x 20 sec on vortex mixer and left to stand for 1 hour to allow the formed air bubbles to clear. Each vial was then counted for 10 cycles of 10 min. Channel A was set for ^{3}H (0-25 keV) and channel B for the remaining ^{14}C signal (25-400 keV). In each sample set, a ^{14}C standard was included to correct for channel overlap between the two isotopes.

8.3.8.5 Processing of the scintillation data:

The isotope effects were calculated from the average cpm/isotope by equation 3.9 (Section 3.1.6). Processing was done in Microsoft Excel 2010. The specific counts for ^{3}H and ^{14}C were found by correcting for channel overlap of ^{14}C. The correction factor $C_{overlap}$ was found through Equation 8.1 from the average CPM of channel A (0-25 keV)($CPMA_{average}$) and the average CPM of channel B (25-400 keV)($CPMB_{average}$) of 10 cycles. with an example given in Table 8.12.

$$\mathrm{CPMA}_{\mathrm{average}}/\mathrm{CPMB}_{\mathrm{average}} = C_{overlap} \qquad \text{Equation 8.1}$$

Table 8.12. An example of the channel overlap factor

$CPMA_{average}$	$CPMB_{average}$	$C_{overlap}$
15345	35620	0.430802

From this factor the proportion of ^{3}H could be found by Equation 8.2 and ^{14}C by equation 8.3 and an example of each is given in table 8.13.

$$CPMA_{average} - \left(CPMB_{avarage} * C_{overlap}\right) = CPM\ ^{3}H \qquad \text{Equation 8.2}$$

$$CPMA_{average} + \left(CPMB_{avarage} * C_{overlap}\right) = CPM\ ^{14}C \qquad \text{Equation 8.3}$$

Table 8.13. An example of the amount of ^{3}H and ^{14}C found from the average CPMs.

$CPMA_{average}$	$CPMB_{average}$	^{3}H	^{14}C
20834	7995	17390	11439

The ^{14}C and ^{3}H ratios for the partial and full conversion experiments could then be found according by equation 8.4 if the internal standard contained ^{3}H and equations 8.5 if the internal standard contained ^{14}C.

$$CPM\ ^{14}C/CPM\ ^{3}H = R_p \qquad \text{Equation 8.4}$$

$$CPM\ ^{3}H/CPM\ ^{14}C = R_p \qquad \text{Equation 8.5}$$

This information could then be used to calculate the ratio of ^{14}C and ^{3}H of the samples and the associated KIEs could be found through Equation 3.9 (Section 3.1.6). fractional conversions were also calculated to ensure partially completed reactions were within the desired range (0.15 to 0.45) (Table 8.14).

Table 8.14. example of a data set and the calculated kinetic isotope effects

	^{3}H	^{14}C	R_p	F	KIE
KIE1	12277	12570	0.977	0.281	1.0029
KIE2	11452	11777	0.972	0.263	1.0080
KIE3	14039	14442	0.972	0.323	1.0088
KIE4	14441	14827	0.974	0.332	1.0064
	^{3}H	**^{14}C**	**R_o**	**Average R_o**	**Average KIE**
100 %	43891	44744	0.981	0.979	1.0065
100 %	43628	44648	0.977		

8.3.9 Measuring commitment to catalysis

The forward commitment to catalysis or the partitioning of the Michaelis complex to product relative to substrate release (k_{cat}/k_{off}) was measured by the isotope trapping method.[85] The experiment determines the fraction of enzyme bound isotopically labelled PRATP that converts to product after a large excess of unlabelled PRATP and the second substrate PA is added. Fractions are taken after one and several turnovers and extrapolated back to find the initial amount of substrate turned over. The commitment factor can be found through equation Equation 12-14 including an example calculation for *C. jejuni* ATP-PRTase Section 3.3.2.

$$\frac{[E_t][^{14}\mathrm{CPRATP}]}{K_\mathrm{m} + [^{14}\mathrm{CPRATP}]} = [EA] \qquad \text{Equation 12}$$

$$\frac{[\mathrm{PRPA}]}{[EA]} = Y \qquad \text{Equation 13}$$

$$\frac{Y}{1-Y} = C_f \qquad \text{Equation 14}$$

***C. jejuni* ATP-PRTase**

$$\frac{[20\ \mu\mathrm{M}][50\ \mu\mathrm{M}]}{10\ \mu\mathrm{M} + [50\ \mu\mathrm{M}]} = [16.6\mu\mathrm{M}]$$

$$\frac{[0.48\ \mu\mathrm{M}]}{[16.6\mu\mathrm{M}]} = 0.029$$

***L. lactis* ATP-PRTase**

$$\frac{[20\ \mu\mathrm{M}][50\ \mu\mathrm{M}]}{2.6\ \mu\mathrm{M} + [50\ \mu\mathrm{M}]} = [19\mu\mathrm{M}]$$

$$\frac{[0.53\ \mu\mathrm{M}]}{[19\mu\mathrm{M}]} = 0.028$$

$$\frac{0.029}{1-0.029} = 0.030 \qquad \frac{0.028}{1-0.028} = 0.029$$

More specifically, enzyme 20 µm (final concentration) was mixed with 50 µM final concentration of [5'-^{14}C]PRATP to give a total volume of 20µL and incubated for 10 seconds before a chase solution containing 500 µM unlabelled PRATP and 10 mM PA was added (480 µL). 120 µL samples were then taken after 1, 2, 4, and 8 catalytic turnovers and quenched onto 10 µL H_2SO_4 0.1 M. samples where processed according to the standard method Section 8.3.8.4. A no enzyme control was added to correct for non-enzymatic degradation. The concentration of formed PRPA was back extrapolated and the y intercept gave the [PRPA] at time 0. The concentratrion of PRPA was calculated based on the fraction of radiation recovered and the total amount of radiation in cpm available. Graphs obtained after 1 and several turnovers can be found in appendix II.

8.4 Experimental procedures for chapter 4

8.4.1 Generation of electrostatic potential maps

The molecular electrostatic potential surfaces (ESP) were calculated from the optimised structures using the Cubegen utility provided with Gaussian 09 program.[113] The electrostatic potential was calculated first, followed by the generation of the ESP map. Maps were generated at an isovalue of 0.05 and the arbitrary colour scale adjusted to represent appropriate colour coding.

8.4.2 ISOEFF98 input file:

The ISOEFF98 program is available free of charge by contacting Prof Piotr Paneth on piotr.paneth@p.lodz.pl

An example of the used input file system is given below. All files were generated with the jEdit text editing program. Highlighted in red requires the input of the mass of the used isotope. Highlighter in yellow are the areas that need referencing to, or input from, the structure. The input file, ISOEFF98.exe, loopfile, and .log file were placed in the same folder. The isoeff98 program was operated from the Cygwin64 terminal by executing the loop file.

The input file, {filename.inp}

```
----1_3H-----
-$ISOEFF HMA(methyl) = Atomic mass of used isotopic substitution (example 3H = 3) KIE= .T. EIE= .F.
$END
 $ISOEFF HMASS(Position label in GS) = 3 HMATS(position label in TS) = 3 KIE= .T. EIE= .F. RMTRIV =
.F.
  cutlts = 1 cutlss = 1 cututs = 10000 cutuss = 10000
  SCFACT=Scaling factor for functional TKELV(1) = Experimental temperature in K $END
```

Loop file:

```
for n in [insert TS file name] *.log; do
        m=`echo $n | cut -d'.' -f1`;
        for x in [insert TS file name]*.log ; do
                y=`echo $x | cut -d'.' -f1`;
                echo \ \$GLOBAL MEMORY=10000000 \$END > $y$m;
                echo \ \$IODEF SSINP = $x TSINP = $n \$END >> $y$m;
                cat The input file.inp >> $y$m;
                ./isoeff.exe $y$m;
                grep "ISOTOPE EFFECT" $y$m.iso > KIE$y$m;
        done;
done
```

8.4.3 Coordinates for the ground state structure

Visualisation of the structures was done using the Gausview program.

Input structure				**Output structure**			
-2 1				-2 1			
O	-9.90400000	-25.26500000	-32.83600000	O	1.74502200	-1.85686900	1.18272500
O	-9.45000000	-26.38500000	-34.75500000	O	3.03877600	-3.15982500	-0.68110200
O	-8.84000000	-27.45000000	-32.60200000	O	3.97982900	-0.89795000	0.25482700
P	-8.91400000	-26.16000000	-33.33200000	P	2.77551700	-1.82832600	0.01124000
C	-6.33100000	-26.47600000	-32.93300000	C	2.30421000	0.23014800	-1.80901700
O	-7.43200000	-25.67900000	-33.33900000	O	1.78321400	-0.88495500	-1.15478300
C	-5.14900000	-25.74100000	-32.25400000	C	1.87397900	1.56433400	-1.17199800
O	-4.49800000	-24.74900000	-33.06800000	O	0.41440300	1.70160100	-1.16503600
C	-5.54500000	-25.02600000	-30.98100000	C	2.27917000	1.71760700	0.28858100
O	-4.56600000	-25.28300000	-30.08400000	O	2.34264400	3.12176500	0.63927300
C	-5.44300000	-23.55400000	-31.36500000	C	1.10773600	1.03891400	1.00427000
O	-5.05300000	-22.85700000	-30.24000000	O	0.93751500	1.50816800	2.35872700
C	-4.29200000	-23.58200000	-32.30700000	C	-0.08819100	1.44253500	0.13513300
N	-4.24600000	-22.54100000	-33.35600000	N	-1.23045400	0.49861300	0.07352400
C	-5.41140028	-22.02535755	-33.83944292	C	-1.07678900	-0.83177100	0.40666400
N	-5.36734329	-21.03058923	-34.78685623	N	-2.06884700	-1.69040500	0.50625800
C	-4.21956770	-20.57246939	-35.23156220	C	-3.27675100	-1.14073200	0.23978900
C	-3.09918645	-21.08912803	-34.75352799	C	-3.57299300	0.15220000	-0.16264700
C	-3.11509181	-22.05158279	-33.83185025	C	-2.47675700	1.07626500	-0.37147500
N	-1.95676789	-22.51382777	-33.41546428	N	-2.61676700	2.24779200	-0.88987900

N	-2.12880867	-20.45224245	-35.35453440	N	-4.93600300	0.30456400	-0.36240700
C	-2.63079301	-19.56887076	-36.19962149	C	-5.44931000	-0.87131200	-0.07734200
N	-3.93736517	-19.65726639	-36.14555301	N	-4.49030500	-1.79279000	0.29488200
H	-6.72123487	-27.13180550	-32.18297607	H	3.40653400	0.21789000	-1.81876300
H	-5.92757125	-26.87875611	-33.83850139	H	1.94088900	0.24127400	-2.85367700
H	-4.47525692	-26.54991012	-32.06259811	H	2.25777000	2.40175300	-1.77088600
H	-6.50213937	-25.30134529	-30.58988722	H	3.22149000	1.20174400	0.50313900
H	-4.71405911	-24.76277211	-29.29087867	H	2.10391800	3.14866400	1.57936200
H	-6.33877825	-23.12029433	-31.75791311	H	1.30092500	-0.04317200	1.02694600
H	-3.41697804	-23.47789620	-31.70004122	H	-0.52124400	2.36479700	0.54611000
H	-6.35204993	-22.39166313	-33.48462121	H	-0.06043600	-1.20049500	0.60878600
H	-1.92245745	-23.17484957	-32.66588242	H	-1.71522400	2.69450200	-1.04967500
H	-2.06991733	-18.89996007	-36.81839201	H	-6.50111900	-1.12294800	-0.11393000
H	-4.59392776	-19.13460816	-36.68938565	H	-4.61202100	-2.76483800	0.53505100
H	-5.09755333	-21.89808667	-30.22998882	H	1.26534900	0.78670800	2.91349100

8.4.4 Coordinates for the truncated transition state input structure

Truncated input structure			
2 1			
Mg	-0.33811700	1.35735400	0.41011000
P	1.48097400	-1.39931200	0.55831100
C	-0.91109400	-1.36040600	-1.02761500
O	0.09358200	-0.59152900	-0.04809800
O	1.27907600	-1.33673200	2.05616200
O	1.43278000	1.82082700	-0.53589900
C	-1.86561000	-0.23125700	-1.46355200
O	-1.26342600	1.05836900	-1.57553800
O	1.73414700	-2.58599400	-0.29371500
C	-2.87155100	-0.20553900	-0.29347300
O	-2.34792800	0.72054300	0.69452100
C	-2.86226600	-1.66615500	0.17290200
O	-1.67012100	-2.27290700	-0.37509300
O	-0.70014700	3.38825700	0.55146600
O	0.06109900	0.91605000	2.35670600
H	-0.26422600	-1.84821300	-1.75239800
H	-2.37444100	-0.50825100	-2.39312800
H	-3.86249300	0.14433100	-0.58984900
H	-2.84613500	-1.76650100	1.26370600
H	-0.06729500	3.99215000	0.12790500
H	0.53848500	0.00140600	2.41238100
H	-0.63533100	1.07335100	-2.31848100
H	-1.57192500	3.81385500	0.52238300
H	0.41478300	1.48536900	3.05624400
C	2.77583800	-0.03566200	0.17210500
H	3.68258700	-0.55010000	-0.15574200
H	2.97630900	0.48471300	1.11004000
C	2.25701200	0.93351300	-0.82135700
O	2.59478300	0.81925600	-2.10061900
H	-3.72129800	-2.22035500	-0.21461200
H	3.24977600	0.10942400	-2.23160800
H	-2.52074600	0.39415400	1.59397700
H	2.68059288	-2.71657153	-0.38746838

8.4.5 Coordinates for the *M. tuberculosis* transitions state structure

Input structure				Output structure			
1 1				1 1			
Mg	-2.23604661	-0.94868445	0.64844343	Mg	-2.21702500	-0.93467800	0.64720600
P	-2.04052161	1.90943555	-0.00376557	P	-1.95099000	1.90296500	-0.03929100
C	0.34428461	-0.77811555	-2.03924043	C	0.32669600	-0.90375800	-2.05065300
O	-1.29630761	0.55931255	-0.28360857	O	-1.23881500	0.53432000	-0.30605900
O	-2.11501261	2.20487055	1.48851143	O	-2.02274400	2.21624600	1.45011100
O	-4.01749561	-0.40527745	0.08938643	O	-3.98444200	-0.36060700	0.07374400
C	-0.05735039	-2.05946055	-1.35631743	C	-0.10106400	-2.13965300	-1.30291600
O	-1.44613139	-2.14812255	-1.19593043	O	-1.49136700	-2.17977000	-1.13556200
O	-1.20322261	3.05151855	-0.80159257	O	-1.08512800	3.01866800	-0.84417800
C	0.75205461	-1.99921455	-0.01096343	C	0.71492600	-2.03265000	0.03675800
O	-0.19135739	-1.78484155	1.02620657	O	-0.22600500	-1.79309500	1.06985100
C	1.76201061	-0.85474455	-0.23324443	C	1.71735600	-0.88822600	-0.22119300
O	1.36227661	-0.23159355	-1.55125243	O	1.34001000	-0.33755000	-1.57934400
O	-3.41054261	-2.71478745	0.87648743	O	-3.43714700	-2.67443700	0.88657200
O	-1.91213261	-0.11790645	2.53307843	O	-1.88774700	-0.09447100	2.52407600
H	-0.08847739	-0.35407855	-2.94430743	H	-0.09421600	-0.52362900	-2.98094100
H	0.32555061	-2.87073855	-1.99609543	H	0.25633200	-2.99202000	-1.90248800
H	1.27498861	-2.94486655	0.16920457	H	1.24485400	-2.96945800	0.24144400
H	1.64298061	-0.02030655	0.46213257	H	1.56805700	-0.02054900	0.42599000
H	-4.20451061	-2.18568845	0.60972643	H	-4.21211900	-2.12363700	0.60575200
H	-1.95888700	0.86998800	2.26932200	H	-1.92075100	0.90438600	2.26169600
H	-1.76910939	-3.06220955	-1.23775843	H	-1.84150800	-3.08500000	-1.12678400
H	-3.65708561	-3.30915045	1.59630543	H	-3.70819900	-3.24639400	1.61576000
H	-2.48038461	-0.25200545	3.30299943	H	-2.46905100	-0.20303100	3.28823300
C	-3.67282061	1.84249855	-0.80906657	C	-3.58347600	1.86843800	-0.84737800
H	-3.51311761	1.81812555	-1.89376457	H	-3.42245800	1.83207500	-1.93151200
H	-4.21303561	2.76903455	-0.59470457	H	-4.10210400	2.80908300	-0.64123900
C	-4.59606361	0.64643055	-0.44293157	C	-4.53576200	0.69783400	-0.47457600
O	-5.78306461	0.73395755	-0.69749457	O	-5.71887800	0.80859400	-0.73793600
H	0.16366161	-1.30533855	1.79248657	H	0.14411900	-1.33752200	1.84343600
H	-1.10690261	3.86745755	-0.28761357	H	-0.98100700	3.84026000	-0.34088000
C	3.22273461	-1.27582155	-0.33726143	C	3.18487900	-1.29342100	-0.27276600
H	3.54890661	-1.60506955	0.65609457	H	3.49298400	-1.56760400	0.74295000
H	3.35056761	-2.10467555	-1.04276943	H	3.33908100	-2.15475500	-0.93260600
O	3.98923761	-0.17969455	-0.81314843	O	3.94773100	-0.21179000	-0.78608900
P	5.02002561	0.64232445	0.13784357	P	4.94686300	0.67076100	0.14442900
O	4.01055861	0.95539745	1.36202157	O	3.90851800	1.02583900	1.33216000
H	4.37665061	1.54702045	2.03809557	H	4.25117600	1.65604000	1.98524500
O	5.97134261	-0.49328955	0.78001157	O	5.90161800	-0.41805500	0.85864300
H	6.87567661	-0.45879355	0.42954357	H	6.81200100	-0.38675800	0.52389900
O	5.69092461	1.75271645	-0.55103143	O	5.61535200	1.75703000	-0.58413800
B 3 4 F							
D 13 12 32 33 F							
D 12 32 35 36 F							

8.4.6 Coordinates for the *C. jejuni* transitions state structure

Input structure	Output structure

1 1			
Mg	-2.23405500	-1.01722900	0.59965900
P	-2.10870800	1.88054300	0.02718500
C	0.33118486	-0.51197748	-1.97387525
O	-1.36624294	0.56967453	-0.25013867
O	-2.18148800	2.13541500	1.52486500
O	-4.02659700	-0.48392600	0.06818900
C	-0.04770700	-1.86053200	-1.41357000
O	-1.42978100	-2.04199100	-1.27041600
O	-1.29610000	3.05267400	-0.75165600
C	0.75711900	-1.91617300	-0.06916200
O	-0.18597400	-1.77037600	0.98276900
C	1.77899400	-0.77051300	-0.20892100
O	1.33587500	0.00142600	-1.42578600
O	-3.36153900	-2.80547400	0.78748200
O	-1.92052000	-0.22179300	2.50559300
H	-0.15409300	0.02212900	-2.78656500
H	0.37606900	-2.57288500	-2.13956500
H	1.26778800	-2.87881600	0.04199800
H	1.71766000	-0.02115700	0.58400600
H	-4.17852800	-2.30249800	0.54689900
H	-2.00745700	0.78019400	2.29268300
H	-1.70160000	-2.96025700	-1.42352200
H	-3.57607200	-3.44765300	1.47554400
H	-2.47801500	-0.39505600	3.27566800
C	-3.73916500	1.79248100	-0.77702400
H	-3.58191700	1.79663900	-1.86234900
H	-4.29988100	2.70117800	-0.54000600
C	-4.63224800	0.56595500	-0.43446500
O	-5.82229300	0.63402400	-0.68060500
H	0.14965200	-1.28156200	1.75180200
H	-1.19020900	3.85237300	-0.21436900
C	3.22390100	-1.20639300	-0.42535600
H	3.58671800	-1.65067700	0.50842200
H	3.30046500	-1.95322100	-1.22377800
O	3.99523800	-0.07917800	-0.81353500
P	5.08572800	0.61333000	0.17219400
O	4.13796900	0.81539600	1.46749800
H	4.55226900	1.31606600	2.18777100
O	6.03395100	-0.60413000	0.64816900
H	6.92353400	-0.55102400	0.26392200
O	5.75630800	1.77683600	-0.42316100
B 3 4 F			
D 13 12 32 33 F			
D 12 32 35 36 F			

1 1			
Mg	-2.22516700	-0.98172300	0.62075600
P	-2.06786400	1.89143200	-0.00273200
C	0.32295500	-0.65955000	-1.99634200
O	-1.30665300	0.55161700	-0.29341700
O	-2.14308800	2.17184000	1.49174700
O	-4.01326700	-0.45020900	0.07187000
C	-0.05813700	-1.97539300	-1.36901800
O	-1.44484200	-2.10031700	-1.21444100
O	-1.24528100	3.04886400	-0.79347600
C	0.74845800	-1.95657200	-0.02200200
O	-0.19542700	-1.76553200	1.02015900
C	1.76373500	-0.81186700	-0.21447500
O	1.34698500	-0.13184000	-1.49692500
O	-3.37488600	-2.76087600	0.83565400
O	-1.90579600	-0.16203000	2.51256500
H	-0.11873800	-0.20072000	-2.87939300
H	0.34188000	-2.75130300	-2.04146600
H	1.26619800	-2.90918000	0.13460000
H	1.66749400	-0.00704000	0.51832900
H	-4.17964700	-2.24431700	0.57892800
H	-1.98224400	0.83874600	2.27717500
H	-1.74778600	-3.01813500	-1.29835200
H	-3.60690600	-3.37285100	1.54538800
H	-2.46750600	-0.31297900	3.28420000
C	-3.69928100	1.80960100	-0.80738800
H	-3.54024500	1.79537200	-1.89237400
H	-4.25085700	2.72763900	-0.58552000
C	-4.60650800	0.59874100	-0.44881300
O	-5.79527400	0.67446600	-0.69890700
H	0.15033400	-1.27639100	1.78458900
H	-1.14839800	3.85868800	-0.26994900
C	3.21870600	-1.24152700	-0.36059700
H	3.55881700	-1.61440000	0.61235300
H	3.32670700	-2.04152600	-1.10181800
O	3.98743600	-0.13358100	-0.80468700
P	5.04168400	0.63975700	0.16100300
O	4.05517800	0.91449300	1.41298800
H	4.43953500	1.47324800	2.10665700
O	5.99099200	-0.52961000	0.74288600
H	6.89009900	-0.49018800	0.37971800
O	5.71347300	1.76991200	-0.49406500

8.4.7 Coordinates for the *L. lactis* transitions state structure

Input structure			
1 1			
Mg	-2.22036400	-1.04975100	0.58013300
P	-2.13281900	1.86779200	0.05837800
C	0.27635122	-0.30495453	-1.88561258
O	-1.34762043	0.56567764	-0.19602219

Output structure			
1 1			
Mg	-2.23755700	-1.03361100	0.58973700
P	-2.11840600	1.87546000	0.03927400
C	0.28473100	-0.41947200	-1.91359300
O	-1.34217400	0.54526500	-0.27882700

O	-2.23933600	2.10971200	1.54945700	O	-2.18798000	2.11995300	1.53803900
O	-4.02026800	-0.52180500	0.08877700	O	-4.03060300	-0.49388800	0.06711800
C	-0.06408900	-1.71196500	-1.46380300	C	-0.04696000	-1.81029300	-1.43391200
O	-1.44085100	-1.94458700	-1.33801300	O	-1.42657700	-2.01589600	-1.29655800
O	-1.32931600	3.05160100	-0.70945700	O	-1.30673400	3.05034400	-0.73613700
C	0.75200100	-1.87444100	-0.13705800	C	0.75826000	-1.90187400	-0.09294400
O	-0.17638200	-1.80712900	0.93871900	O	-0.18391500	-1.77857600	0.96368300
C	1.77744700	-0.72735900	-0.20618300	C	1.78215400	-0.75561900	-0.20713800
O	1.30786200	0.15188600	-1.32852000	O	1.32813800	0.05654600	-1.39178300
O	-3.29693700	-2.86631700	0.67212400	O	-3.36012400	-2.82220600	0.76661500
O	-1.90534700	-0.30043700	2.50019700	O	-1.92510700	-0.24903100	2.50162300
H	-0.20810200	0.28732000	-2.65471700	H	-0.17162000	0.11754100	-2.74194300
H	0.37351000	-2.33912700	-2.25723100	H	0.38565200	-2.49200900	-2.18357900
H	1.25648600	-2.84572200	-0.10082600	H	1.26613000	-2.86829100	-0.00460000
H	1.75781900	-0.05787400	0.65690700	H	1.73491800	-0.03394900	0.61202300
H	-4.10553200	-2.33671500	0.43830300	H	-4.18151400	-2.32384100	0.53366900
H	-2.00631700	0.69902900	2.30723900	H	-2.01262600	0.75347800	2.29858200
H	-1.65935500	-2.88485400	-1.46455400	H	-1.68492300	-2.93223800	-1.48045200
H	-3.52798500	-3.44133700	1.41791800	H	-3.56768800	-3.47782400	1.44398400
H	-2.50881000	-0.48836400	3.23686900	H	-2.48247200	-0.42970600	3.27013200
C	-3.73107100	1.73457600	-0.79525100	C	-3.74917200	1.78858300	-0.76357700
H	-3.53542900	1.67150500	-1.87349900	H	-3.59339200	1.79859300	-1.84907900
H	-4.29943400	2.65536800	-0.63559100	H	-4.31088300	2.69516900	-0.52089300
C	-4.63831900	0.53070100	-0.40228100	C	-4.63946400	0.55844400	-0.42558400
O	-5.83104300	0.61135300	-0.60215900	O	-5.83049500	0.62744600	-0.66658800
H	0.13725900	-1.27212900	1.69005500	H	0.14982900	-1.29647800	1.73767100
H	-1.18552100	3.81483000	-0.12416400	H	-1.19424300	3.84613200	-0.19433400
C	3.21032900	-1.15454800	-0.51033300	C	3.22258200	-1.19054300	-0.45464600
H	3.60134000	-1.68551200	0.36472300	H	3.59510700	-1.66646600	0.45943500
H	3.24476400	-1.82965000	-1.37329800	H	3.28616400	-1.91113400	-1.27791900
O	3.97531800	-0.00434900	-0.82716500	O	3.99460300	-0.05463400	-0.81572100
P	5.10180100	0.58080800	0.18795800	P	5.09969500	0.60140100	0.17816400
O	4.19167400	0.66833300	1.52283100	O	4.16812100	0.76856800	1.49020300
H	4.66546200	1.06912100	2.27273800	H	4.59455600	1.24319700	2.22094500
O	6.04901200	-0.68167800	0.52786900	O	6.04683700	-0.63480600	0.60589300
H	6.91355600	-0.60016400	0.08759700	H	6.93237600	-0.57405600	0.21354100
O	5.76761800	1.78462200	-0.32043700	O	5.77016000	1.77928400	-0.38846600

B 3 4 F
D 13 12 32 33 F
D 12 32 35 36 F

8.5 Experimental procedures for chapter 5

8.5.1 Lanzetta assay for detection and quantification of phosphates

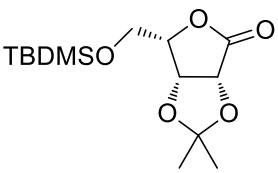
The Lanzetta reagent was made according to published procedures.[206] The reagent was made fresh prior to use by combining 3 parts 0.045% w/v malachite green in water, 1 part 4.2% w/v ammonium molybdate in 4M HCl and 0.1 part 1.5% v/v Triton X-100 in water. The parts were combined in a beaker wrapped in tin foil and stirred in the dark for at least 1 hour. The solution was then filtered through a 0.45-micron syringe filter again avoiding exposure to direct light as much as possible.

Calabration curve: KH_2PO_4 was dried at 100 °C for at least one day and followed by further drying under vacuum overnight in a desiccator. A calibration curve for the determination of phosphate concentration was made ranging from 9 µM to 150 µM. 10 µL of stock solution was mixed with 700 µL of Lanzetta reagent and samples incubate for 20 min in the dark prior to measure the absorbance at 600 nm .

The concentration of inhibitor was measured by cleaving the phosphate group with calf alkaline phosphatase (5 units/mL in 4mM $MgCl_2$). Sample preparation: 40 µL of inhibitor solution (within 25-100 µM based on weight), 10 µL alkaline phosphatase stock, and 50 µL Tris buffer (100 mM NaCl, 5 mM $MgCl_2$, and 50 mM Tris HCl at pH 8.0) were incubated overnight. 20 µL of digested sample was mixed with 700 µL Lanzetta reagent and then incubate for 20 min in the dark prior to measure the absorbance at 600 nm. Data was fitted again the calibration and the stock concentration back calculated. Measurements were carried out in duplicates or triplicates.

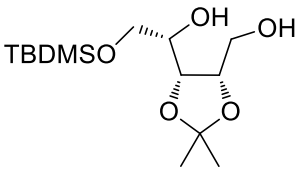
8.5.2 Synthesis of transition state analogues

8.5.2.1 Synthesis of the iminoribitol based transition state analogues



5-*O*-*tert*-butyldimethylsilyl-2,3-*O*-isopropylidene-L-lyxono lactone (5.8). The isopropylidene lactone 5.7 (5 g, 26.6 mmol), was dissolved in dry DMF (45 mL) cooled to 0°C followed by the addition of imidazole (2.7 g, 40 mmol) and tert-butyldimethylsilyl chloride (6 g, 40 mmol). The solution was allowed to warm up to room temperature (20°C). After 4 h, TLC analysis (PE/EtOAc, 4/1, v/v) indicated complete conversion of starting material (R_f 0.12) into a major product (R_f 0.52). The reaction mixture was concentrated. The residue was then purified by flash chromatography (EtOAc– PE, 1 : 6) to give the fully protected lactone 5.8 (5.47 g, 68%) as a white crystalline solid, **Crystallography:** Single crystals of $C_{14}H_{26}O_5Si$ were grown by evaporation from Petroleum ether and ethyl acetate. A suitable crystal was selected and mounted on a nylon loop in perfluoronated oil on a SuperNova, Dual, Cu at zero, Atlas diffractometer. The crystal was kept at 120.01(10) K during data collection. Using Olex2, the structure was solved with the ShelXS structure solution program using direct methods and refined with the ShelXL refinement package using Least Squares minimisation.[207–209] refinement parameters can be found in appendix IV. mp 90–91 ◦C R_f :

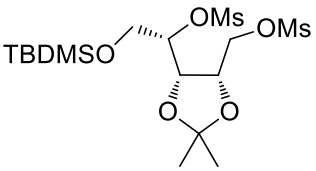
0.52 (Pe/EtOAc, 1/1, v/v): IR (ν): 462.1, 506.3, 723.8, 778.9, 843.4, 873.1, 917.1, 975.2, 1034.0, 1092.6, 1126.5, 1176.1, 1225.5, 1265.3, 1774.6, 1784.4 cm^{-1}): ^{1}H NMR (400 MHz, $CDCl_3$); δ 0.10 (s, 6H, 2x Me OTBDMS), 0.90 (s, 9H, t-Bu OTBDMS), 1.39 (s, 3H, Me iso-prop), 1.46 (s, 3H, Me iso-prop), 3.96 (ddd, *J* = 6.7, 6.2 Hz, 2H, H5), 4.52 (dt, *J* = 6.6, 2.7 Hz, 1 H, H4), 4.77 - 4.87 (m, 2H, H2, H3); ^{13}C NMR (101 MHz, $CDCl_3$); δ -5.38 (Me OTBDMS), -5.21 (Me OTBDMS), 18.45 (Cq, t-Bu OTBDMS), 25.93 (Me, t-Bu OTBDMS), 26.02 (Me iso-prop), 26.92 (Me iso-prop), 61.02 (C5), 75.89, 76.19 (C2, C3), 79.53 (C4), 114.20 (Cq, iso-prop), 173.86 (C1). HRMS(ESI) *m/z* calcld. for $[C_{16}H_{35}O_9S_2Si]^+$: 303.1622, obsd.: 303.1619. **Crystal Data** for $C_{14}H_{26}O_5Si$ (*M* =302.44 g/mol): monoclinic, space group $P2_1$ (no. 4), *a* = 5.95513(12) Å, *b* = 10.14114(19) Å, *c* = 13.3855(3) Å, *β* = 98.4789(19)°, *V* = 799.54(3) $Å^3$, *Z* = 2, *T* = 120.01(10) K, μ(Cu Kα) = 1.442 mm^{-1}, *Dcalc* = 1.256 g/cm^3, 7139 reflections measured (13.376° ≤ 2Θ ≤ 147.52°), 3067 unique (R_{int} = 0.0193, R_{sigma} = 0.0195) which were used in all calculations. The final R_1 was 0.0253 (I > 2σ(I)) and wR_2 was 0.0690 (all data).



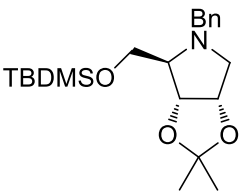
5-*O*-*tert*-butyldimethylsilyl-2,3-*O*-isopropylidene-L-lyxol (5.6).

Protected lactone **5.8** (3.9g, 12.9 mmol) was dissolved in MeOH (24 mL), cooled to 0°C and sodium borohydride (535 mg, 14.2 mmol). The solution was allowed to warm up to rt. After 7 hours TLC analysis (PE/EtOAc, 2/1, v/v) showed residual starting material as well as a major product (R_f 0.23). More sodium borohydride (530 mg, 14.2 mmol) was added at O°C. After a further 16 h, TLC showed complete consumption of starting material. The reaction was quenched by the addition of water, diluted with EtOAc and the organic phase washed with H_2O (2x), and brine (1x), dried over $MgSO_4$ filtered and concentrated *in vacuo*. The crude reaction mixture was used without further purification in the next reaction (3.87 g, 12.6). R_f : 0.23 (Pe/EtOAc, 2/1, v/v): IR (ν): 474.8, 511.0, 556.9, 617.8, 668.6, 722.2, 746.1, 772.2, 836.8, 894.9, 1045.6, 1059.5, 1081.5, 1105.6, 1133.6, 1168.9, 1212.5, 1257.2, 3229.4 cm^{-1}): ^{1}H NMR (400 MHz, CD_3OD); δ 0.00 (s, 3H, Me OTBDMS), 0.01 (s, 3H, Me OTBDMS), 0.83 (s, 9H, t-Bu OTBDMS), 1.23 (s, 3H, Me iso-prop), 1.39 (s, 3H, Me iso-prop), 3.57 - 3.60 (m, 2H, H1), 3.62 - 3.72 (m, 3H, C1, H4), 4.10 - 4.16 (m, 1H, H3), 4.2 (dd, 1H, *J* = 2.8 Hz, *J* = 7.1 Hz, H3); ^{13}C NMR (101 MHz, CD_3OD); δ -6.74 (Me OTBDMS), -6.67 (Me OTBDMS), 17.80 (Cq, *t*-Bu OTBDMS), 23.96

(Me iso-prop), 24.97 (Me, *t*-Bu OTBDMS), 25.95 (Me iso-prop), 60.82 (C5), 64.32 (C1), 69.13 (C4), 75.73 (C3), 77.65 (C2), 107.78 (Cq, iso-prop); HRMS(ESI) *m/z* calcld. for $[C_{14}H_{27}O_5Si]^+$: 307.1935, obsd.: 307.1934

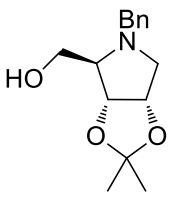


5-*O*-*tert*-butyldimethylsilyl-2,3-*O*-isopropylidene-1,4-*O*-di-methansulfonl-L-lyxol (5.9). To a cooled (0°C) solution of DMAP (402 mg, 3.3 mmol) in dry pyridine (30 mL), methanesulphonyl chloride (1.25 mL, 16.3 mmol) was added and left to stir for 10 min upon which solids are formed. Diol **5.6** (2.1 g, 6.53 mmol)was co-evaporated 2x toluene and added to the reaction in CH_2Cl_2 (20 mL), the resulting mixture left to assume room temperature. After 15 hours TLC analysis (PE/EtOAc, 2/1, v/v) showed complete consumption of starting material and the formation of a higher running product (R_f 0.43). The reaction was diluted with CH_2Cl_2 and poored onto saturated aq. $NaHCO_3$, the aqueous phase was washed with CH_2Cl_2 (3x). The combined organic layers was washed with brine (1x), dried over $MgSO_4$ filtered and concentrated *in vacuo*. Purification by flash chromatography (silica gel, 4/1→2/1 PE/EtOAc) than gave dimesyl **5.9** in a 98% yield (3.0 g, 6.5 mmol) as an amorphous white solid. R_f : 0.43 (Pe/EtOAc, 2/1, v/v): IR (v): 528.3, 722.1, 780.4, 780.4, 837.6, 966.6, 1081.3, 1124.9, 1177.8, 1221.8, 1256.7, cm^{-1}): ^1H NMR (400 MHz, $CDCl_3$); δ 0.09 (s, 3H, Me OTBDMS), 0.10 (s, 3H, Me OTBDMS), 0.90 (s, 9H, t-Bu OTBDMS), 1.38 (s, 3H, Me iso-prop), 1.51 (s, 3H, Me iso-prop), 3.07 (s, 3H, Me OMs), 3.11 (s, 3H, Me OMs), 3.83 (dd, *J* = 10.9, 6.1 Hz, 1H, H5b), 3.95 (dd, J=10.9, 5.4 Hz, 1H, H5a), 4.36 - 4.40 (m, 2H, H1), 4.40 - 4.46 (m, 2H, H2,H3), 4.72-4.77 (m, 1H, H4); ^{13}C NMR (101 MHz, $CDCl_3$); δ -5.38 (2x Me OTBDMS), 18.44 (Cq, t-Bu OTBDMS), 25.93 26.02 (Me iso-prop), 26.92 25.56 (Me iso-prop), 25.98 (Me, t-Bu OTBDMS), 27.39 (Me iso-prop), 37.81 (Me OMs), 39.08 (Me OMs), 63.27 (C5), 68.05 (C1), 74.65, 75.55 (C2, C3), 79.09 (C4), 109.75 (Cq, iso-prop); HRMS(ESI) *m/z* calcld. for $[C_{16}H_{35}O_9S_2Si]^+$: 463.1486, obsd.: 463.1485.



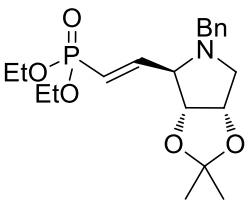
***N*-benzyl-5-*O*-*tert*-butyldimethylsilyl-1,4-dideoxy-1,4-imino-2,3-*O*-isopropylidene-D-ribitol (5.5).** Dimesyl **5.9** (3 g, 6.5 mmol) was taken

up in benzylamine (32 mL) and heated to 160°C with stirring. After 3 hour TLC analysis (PE/EtOAc, 2/1, v/v) showed complete consumption of starting material (R_f 0.43) and the reaction mixture cooled to room temperature. The reaction mixture was subjected to purification by flash chromatography (silica gel, CH_2Cl_2/PE, 1/1, v/v then PE/EtOAc, 10/1, v/v) to give protected iminosugar **5.5** in a 93% yield (2.29 g, 6.06 mmol) as a pale yellow oil. R_f : 0.90 (PE/EtOAc, 2/1, v/v), R_f : 0.30 (CH_2Cl_2/PE, 2/1, v/v): IR (v): 698.2, 722.4, 776.2, 837.6, 1058.6, 1107.4, 1159.8, 1209.5, 1254.5 cm^{-1}): ^{1}H NMR (400 MHz, $CDCl_3$); δ 0.05 (s, 3H, Me OTBDMS), 0.06 (s, 3H, Me OTBDMS), 0.87 - 0.92 (s, 9H, t-Bu OTBDMS), 1.33 (s, 3H, Me iso-prop), 1.56 (s, 3H, Me iso-prop), 2.72 (dd, *J*=10.4, 2.8 Hz, 1H, H1b), 3.01 (dt, *J*=4.21, 2.15 Hz, 1 H), 3.10 (dd, *J*=10.3, 5.6 Hz, 1H, H1a), 3.64 (dd, *J*=10.7, 4.2 Hz, 1H, H5b), 3.71 (d, J=13.5 Hz, 1H, *N*-CH_2 Bn), 3.77 (dd, *J*=10.7, 4.4 Hz, 1H, H5a), 4.01 (d, *J*=13.5 Hz, 1H, *N*-CH_2 Bn), 4.55 (dd, *J*=6.5, 2.1 Hz, 1H, H3), 4.65 (dd, *J*=6.2, 2.9 Hz, 1H, H2), 7.17 - 7.36 (m, 5H H_{arom} Bn); ^{13}C NMR (101 MHz, $CDCl_3$); δ -5.36 (Me OTBDMS), -5.30 (Me OTBDMS), 18.32 (Cq t-Bu OTBDMS), 25.34 (Me iso-prop), 26.05 (Me t-Bu OTBDMS), 27.33 (Me iso-prop), 57.02 (CH_2 Bn), 59.41 (C1), 63.29 (C5), 69.06 (C4), 79.59 (C2), 83.46 (C3), 112.01 (Cq iso-prop), 126.97, 128.38, 128.61, 139.60 (Carom Bn); HRMS(ESI) *m/z* calcld. for $[C_{21}H_{36}NO_3Si]^+$: 378.2459, obsd.: 378.2466.



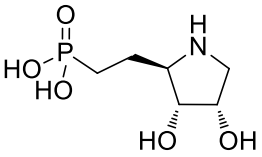
***N*-benzyl-1,4-dideoxy-1,4-imino-2,3-*O*-isopropylidene-D-ribitol (5.4).** To a solution of protected iminosugar **5.5** (1.1 g, 2.9 mmol) in THF (30 mL), tetrabutylammonium fluoride trihydrate (1.4 g, 4.4 mmol) was added at 0°C. The reaction was left to assume room temperature. After 3 hours TLC analysis (PE/EtOAc, 2/1, v/v) showed complete consumption of starting material R_f (0.90) to a lower running product (R_f 0.20). Purification by flash chromatography (silica gel, 3/1→1/1 PE/EtOAc) than gave alcohol **5.4** in a 99% yield (763.1 mg, 2.89 mmol) as a pale yellow oil. R_f : 0.20 (PE/EtOAc, 1/1, v/v): IR (v): 515.1, 699.1, 741.1, 865.5, 1059.1, 1159.2, 1209.8, 1254.5, 1604.0, 1812.9, 1882.7, 1951.8, 3445.0 cm^{-1}): ^{1}H NMR (400 MHz, $CDCl_3$); δ 1.32 (s, 3H, Me iso-prop), 1.53 (s, 3H, Me iso-prop), 2.59 - 2.66 (m, 1H, H1b), 2.93 - 2.98 (m, 1H, H4), 3.15 - 3.24 (m, 1H, H1a), 3.56 (dd, *J*=11.2, 3.7 Hz, 1H, H5b), 3.59 (d, *J*=12.8 Hz, 1H, *N*-CH_2 Bn), 3.69 (dd, *J*=11.2, 3.8 Hz, 1H, H1a), 3.97 (d, *J*= 13.0 Hz, 1H, *N*-CH_2 Bn), 4.55 - 4.63 (m, 2H, H2,H3), 7.23 - 7.36 (m, 5H, Bn Arom); ^{13}C NMR (101 MHz, $CDCl_3$); δ 25.05 (Me iso-prop), 27.44 (Me

iso-prop), 58.36 (*N*-CH_2 Bn), 58.63 (C1), 59.62 (C5), 70.08 (C4), 78.64, 82.71 (C2,C3), 112.90 (Cq iso-prop), 127.50, 128.64, 128.85, 138.58 (Carom Bn); HRMS(ESI) *m/z* calcld. for $[C_{15}H_{22}NO_3]^+$: 264.1594, obsd.: 264.1597.

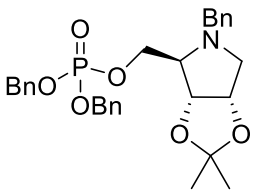


(2*S*,3*R*,4*S*)-1-benzyl-3,4-*O*-isopropylidine-2-carbaldehyde (5.18). To a solution of oxalyl chloride (261 µL, 3.04 mmol) in dry CH_2Cl_2 (15 mL) (-78°C), dimethylsulfoxide (431 µL, 6.08 mmol) was added and left to stir for 60 min at -78°C. Alcohol **5.4** (198.8 g, 0.76 mmol,) was co-evaporated three times with toluene and added dropwise in CH_2Cl_2 (15 mL) and kept at -78°C. After 45 min. triethylamine (1.3 mL, 9.12 mmol) was added. The reaction was kept at -78°C for an additional 2 hours and then allowed to warm up. When TLC analysis showed complete consumption of starting material the reaction was diluted with CH_2Cl_2 saturated $NaHCO_3$ was added and the aqueous layer extracted with CH_2Cl_2 (2x). The combined organic fractions were washed with Brine (1x), dried over $MgSO_4$, filtered and concentrated. The crude was then taken up in toluene and triethylamine salts filtered off. Product formation was confirmed as the hydrate by HRMS(ESI) *m/z* calcld. for $[C_{15}H_{22}NO_4]^+$: 280.1543 obsd.: 280.1543). The aldehyde **5.17** was used without further purification. **diethyl (*E*)-2-((2R,3R,4S)-1-benzyl-3,4-*O*-isopropylidine-pyrrolidin-2-yl)-vinylphosphonate (29).** To a solution of tetraethyl methylenediphosphonate (245 µL, 0.99 mmol) in dry THF (25 mL) at -78°C *n*-BuLi (0.73 mL, 1.42M in THF, 0.99 mmol) was added via the cold wall of the flask. The resulting mixture was allowed to warm up to 0°C. Aldehyde **5.17** was co-evaporated with toluene (3x) and added drupwise to the reaction. The reaction mixture was allowed to assume room temperature. After 14 hours t.l.c analysis (EtOAc), showed formation of a major product (R_f 0.42). The reaction was quenched with saturated aq. NH_4Cl, diluted with CH_2Cl_2 and the aqueous layer was washed with CH_2Cl_2 (2x), the combined organic layers were dried over $MgSO_4$, filtered and concentrated *in vacuo*. Purification by flash chromatography (silica gel, PE/EtOAc 1/1→0/1) than gave vinyl phosphonate **5.18** in a 56% yield (167.6 mg, 0.42 mmol) over two steps as a yellow amorphous solid. R_f : 0.42 (EtOAc): IR (v): 513.6, 722.1, 966.0, 1029.9, 1053.8, 1158.1, 1210.4, 1244.2 cm^{-1}): ^1H NMR (400 MHz, $CDCl_3$); δ 1.30 (s, 3H, Me iso-prop), 1.33 (t, *J*= 7.09 Hz, 3H, Me OEt), 1.34 (t, *J*= 7.20 Hz, 3H, Me OEt), 1.53 (s, 3H, Me iso-prop), 2.58 (dd, *J*= 10.6, 3.6 Hz, 1H, H2b), 3.12 (dd, *J*= 10.6, 5.8

Hz, 1H, H2a), 3.32-3.36 (m, 1H, H5), 3.41 (d, *J*= 13.3 Hz, 1H, *N*-CH_2a Bn), 3.86 (d, *J*= 13.3 Hz, 1H, *N*-CH_2b Bn), 4.00 - 4.16 (m, 4H, CH_2 OEt), 4.41 (dd, *J*= 6.7, 3.8 Hz, 1H, H4), 4.66 (ddd, *J*= 6.2, 3.8 Hz, 1H, H3), 5.97 (dd, *J*= 20.4, 17.2 Hz, 1H, H7), 6.72 (dd, *J*= 17.1, 7.8 Hz, 1H, H6), 7.19 - 7.36 (m, 5H, Bn arom); ^{13}C NMR (101 MHz, $CDCl_3$); δ 16.51 (Me OEt), 16.58 (Me OEt), 25.34 (Me iso-prop), 27.31 (Me iso-prop), 57.19 (CH_2 Bn), 57.97 (C2), 61.93, 62.03 (CH_2 OEt), 71.97 (d, *J*= 22.1 Hz, 1C C5), 78.59 (C3), 84.91 (C4), 113.25 (Cq iso-prop), 120.7 (d, *J*= 186.9 Hz, 1C, C7), 121.64, 127.31, 128.48, 128.66, 128.69, 138.32 (Carom Bn), 150.17, 150.22 (d, *J*= 5.3 Hz, 1C, C6). ^{31}P NMR (162 MHz, $CDCl_3$); δ 17.30 (s, 1P); HRMS(ESI) *m*/*z* calcld. for $[C_{20}H_{31}NO_5P]^+$: 396.1934, obsd.: 396.1944.

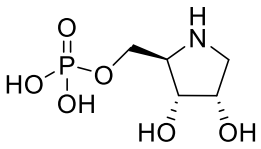


2-([2*R*,3*R*,4*S*]-3,4-dihydroxypyrrolidin-2-yl]ethyl phosphonic acid (5.2). To a solution of vinyl phosphonate **5.19** (75 mg, 0.19 mmol) in CH_2Cl_2, bromotrimethylsylane (131 µL 1.0 mmol) was added. After 3 hours TLC analysis (CH_2Cl_2/MeOH, 9/1, v/v) residual starting material as well as a major product (R_f 0.1). More bromotrimethylsylane (75 µL 0.5 mmol) was added and left for an additional 2 hours after which TLC analysis showed complete consumption of starting material. The reaction was the reaction was concentrated and product formation confirmed by HRMS(ESI) *m*/*z* calcld. for $[C_{16}H_{23}NO_5P]^+$: 340.1308, obsd.: 340.1308. The crude reaction mixture was dissolved in glacial acetic acid (2.5 mL), water (2,5 mL) was added and the resulting mixture was stirred at 70^{o}C for 2 hours. Complete conversion was confirmed by NMR with loss of methyl peaks. The reaction was concentrated and used without further purification. The crude reaction mixture was dissolved in MeOH and hydrogen gas was bubbled through with a balloon for 15 min, 10 mg 10% palladium on carbon was then added and left to stir under balloon pressure. The reaction was stirred under balloon pressure for 48 hours after which complete conversion was confirmed by HRMS. Purification by reverse phase HP20 then gave iminosugar **5.2** in 83% yield over three steps. ^{1}H NMR (400 MHz, D_2O); δ 1.79-1.82 (m, 2H), 1.92-2.17 (m, 2H), 3.31-3.34 (m, 1H), 3.52-3.56 (m, 2H), 4.09-4.12 (m, 1H), 4.38 (bs, 1H); HRMS(ESI) *m*/*z* calcld. for $[C_6H_{15}NO_5P]^+$: 212.0682, obsd.: 212.0686.



***N*-benzyl-6-*O*-dibenzylphosphoryl-1,4-dideoxy-1,4-imino-2,3-*O*-isopropylidene-D-ribitol (5.10).** To a dried solution of **5.4** (107 mg,

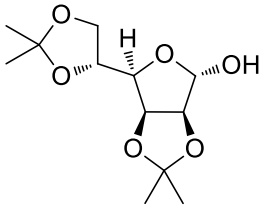
0.38 mmol) in dichloromethane (7.0 mL), tetrazole (85 mg, 1.14 mmol) was added. The resulting suspension was left to stir for 10 min at room temperature before dibenzyl N,N-di-isopropyl-phosphoramidite was added dropwise over 5 min after which most of the tetracole had dissolved. After 1.5 hours TLC analysis showed full conversion of starting material into a higher running product (Pe/EtOAc, 2/1, v/v, R_f 0.15). The reaction mixture was cooled down to -40 °C resulting in a cloudy solution. *m*-CPBA (131 mg, 0.76 mmol) was added in one portion and allowed to reach room temperature. After 2 hours TLC analysis (PE/EtOAc, 1/1, v/v) showed complete consumption of starting material R_f to a lower running product (R_f 0.50). Purification by flash chromatography (silica gel, 4/1→1/1 PE/EtOAc) than gave protected phosphate **5.10** in a 42% yield (80 mg, 0.16 mmol) as a pale yellow oil. R_f : 0.20 (PE/EtOAc, 1/1, v/v): IR (v): 492.9, 594.9, 695.5, 733.8, 860.8, 996.4, 1156.8, 1210.5, 1256.1, 1379.3, 1455.0, 1496.4, 2935.8 cm^{-1}: ^1H NMR (400 MHz, $CDCl_3$); δ 1.25 (s, 3H, Me iso-prop) 1.45 (s, 3H, Me iso-prop), 2.57 (dd, *J* = 3.2 Hz, *J* = 10.4 Hz, 1H, H1*b*), 2.95-2.99 (m, 2H, H1*a*, H4), 3.57 (d, *J*=13.2 Hz, 1H, *N*-CH_2 Bn), 3.86 (d, *J*=13.2 Hz, 1H, *N*-CH_2 Bn), 3.98-4.01 (m, 2H, H5), 4.44-4.46 (m, 1H, H3), 4.50-4.54 (m, 1H, H2), 4.99-5.09 (m, 4H, *O*-CH_2 Bn), 7.18-7.21, 7.23-7.26, 7.35 (m, 15H, Bn Arom); ^{13}C NMR (101 MHz, $CDCl_3$); δ 23.72 (Me iso-prop) 25.91 (Me iso-prop) 56.84 (*N*-CH_2 Bn), 58.00 (C1), 66.09 (C5), 66.78 (C4), 69.52, 69.58 (*O*-CH_2 Bn), 78.53 (C2), 82.36 (C3), 112.30 (Cq iso-prop), 126.90, 127.02, 127.09, 127.76, 127.82, 127.95, 128.31, 128.40, 128.46, 135.69, 135.75 (Carom Bn); HRMS(ESI) *m*/*z* calcld. for $[C_{29}H_{35}NO_6P]^+$: 524.2197, obsd.: 524.2220.



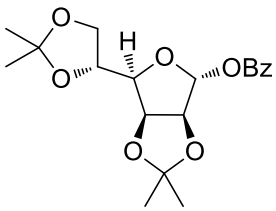
6-*O*-phosphoryl-1,4-dideoxy-1,4-imino-2,3-*O*-isopropylidene-D-ribitol (5.10). Protected phosphate **5.10** (14.0 mg, 0.028mmol) was taken op in methanol (3 mL) and acidified with concentrated HCl (0.2 mL). The resulting solution was left for 5 min and then concentrated to dryness by rotary evaporation. Additional methanol was added and the solvent again removed by rotary evaporation. TLC analysis of the reaction showed a single product (CH_2Cl_2/methanl, 95/5, v/v, R_f 0.3). Longer reaction times resulted in loss off one of the benzyl protecting groups (CH_2Cl_2/methanl, 95/5, v/v, R_f 0.15). Crude product was carried through without further purification. The 2,3-hydroxyl **5.10** was then taken up in ethanol (4 mL) 5 mg 10% palladium on carbon was added and hydrogen gas bubbled through the solution for 15 min and then sealed and left to stir for 6 hours. Purification by reverse phase chromatography (C18, H_2O)

than gave the title compound **5.1** in 67% yield (4 mg, 0.019 mmol) as a pale yellow oil. IR (v): 932.9, 1048.5, 1633.3, 3256.6 cm^{-1}: ^{1}H NMR (400 MHz, D_2O); δ 3.24 (dd, *J*=2.0 Hz, *J*=12.8 Hz, 1H), 3.34-3.38 (m, 1H), 3.58-3.62 (m, 1H), 3.91-3.97 (m, 1H), 4.09-4.15 (m, 1H), 4.21-4.24 (m, 1H), 4.26-4.28 (m, 1H); ^{13}C NMR (101 MHz, $CDCl_3$); δ 49.53, 60.78, 60.84, 69.47, 70.90; ^{31}P NMR (162 MHz, $CDCl_3$); δ 1.61 (s, 1P); HRMS(ESI) *m/z* calcld. for $[C_{12}H_{19}NO_6P]^+$: 304.0945, obsd.: 304.0946.

8.5.2.2 Synthesis of the cyclopentene transition state analogue

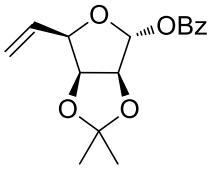


2,3:5,6-di-*O*-isopropylidene-α-D-mannose (5.24). To a suspension of D-mannose **5.23** (10 g, 55.5 mmol) in dry acetone, H_2SO_4 (0.75 mL, 13.9 mmol) was added followed by 2,2-dimethoxy propane (15 mL, 122 mmol). The resulting mixture was stirred until TLC analysis showed complete conversion (R_f 0.40 PE/EtOAc, 1/1, v/v). Triethylamine (4.2 mL, 30 mmol) was added and the resulting mixture concentrated *in vacuo*. Purification by flash chromatography (silica gel, PE/EtOAc 5/1→2/1) than gave iso-propylidene **5.24** in a 79% yield (11.4 g, 43.7 mmol) as a white crystalline solid. R_f 0.40 (PE/EtOAc, 1/1, v/v): IR (v): 514.9, 722.2, 838.9, 975.8, 1062.5, 1166.6, 3431.0 cm^{-1}): R_f : 0.42 (EtOAc): R_f 0.40 (PE/EtOAc, 1/1, v/v): ^{1}H NMR (400 MHz, $CDCl_3$); δ 1.32, 1.37, 1.45, 146 (4x s, 3H, Me iso-prop), 2.90 (bs, 1H, 1-OH), 4.04 (dd, *J*= 8.7, 5.0 Hz, 1H, H6b), 4.08 (dd, *J*= 8.7, 6.2 Hz, 1H, H6a), 4.18 (dd, *J*= 7.1, 3.6 Hz, 1H, H4), 4.40 (dd, *J*= 11.7, 6.2 Hz, 1H, H5), 4.61 (d, *J*= 5.9 Hz, 1H, H2), 4.80 (dd, *J*= 5.9, 3.6 Hz, 1H, H3), 5.37 (s, 1H, H1); ^{13}C NMR (101 MHz, $CDCl_3$); δ 24.63, 25.33, 26.00, 27.00 (4x Me iso-prop), 66.73 (C6), 73.40 (C5), 79.81 (C3), 80.51 (C4), 85.64 (C2), 101.46 (C1), 109.24, 112.82 (2x Cq iso-prop); HRMS(ESI) *m/z* calcld. for $[C_{12}H_{20}NaO_6]^+$: 283.1152, obsd.: 283.1147.



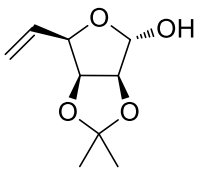
Benzoyl 2,3:5,6-di-*O*-isopropylidene-α-D-mannofuranose (5.25). iso-propylidene **5.24** (6.69 g, 25.7mmol) was co-evaporated with pyridine (1x) under a nitrogen atmosphere, dissolved in dry pyridine (65 mL) and cooled to 0°C. Benzoyl chloride (6 mL) was added drop wise over 2 min followed by *p*-dimethylaminopyridine (628 mg, 5.14 mmol), the resulting

reaction mixture was allowed to assume room temperature. After 2.5 hours t.l.c analysis (PE/EtOAc, 2/1, v/v) showed complete disappearance of starting material and the formation of a higher running product (R_f 0.53). The reaction was poured onto saturated aq. $NaHCO_3$, diluted with EtOAc and the organic layers washed with saturated aq. $NaHCO_3$ (2x), H_2O (1x) and Brine (1x). The organic layer was dried over $MgSO_4$, filtered, and concentrated *in vacuo*. Purification by flash chromatography (silica gel, PE/EtOAc 5/1→2/1) than gave benzoylated product **5.25** in a 86% yield (11.4 g, 43.7 mmol) as a white crystalline solid. R_f 0.53 (PE/EtOAc, 2/1, v/v): IR (v): 516.0, 720.4, 848.4, 969.9, 1035.7, 1072.0, 1084.1, 1208.3, 1255.7, 1292.1, 1724.4 cm^{-1}: ^1H NMR (400 MHz, $CDCl_3$); δ 1.375, 1.383, 1.46, 1.53 (4x s, 3H, Me iso-prop), 4.05 (dd, *J*= 8.9, 4.4 Hz, 1H, H6b), 4.09 - 4.16 (m, 2H, H4, H6a), 4.40 - 4.48 (m, 1H, H5), 4.88 (d, *J*= 5.9 Hz, 1H, H2), 4.94 (dd, *J*= 5.8, 3.6 Hz, 1H, H3), 6.37 (s, 1H), 7.41 - 7.49 (m, 2H, H_{arom} *O*-Bz), 7.56 - 7.62 (m, 1H, H_{arom} *O*-Bz), 7.99 - 8.04 (m, 2H, H_{arom} *O*-Bz); ^{13}C NMR (101 MHz, $CDCl_3$); δ 24.85, 25.27, 26.13 27.12 (4x Me iso-prop), 67.03 (C6), 73.05 (C5), 79.55 (C3), 82.76 (C4), 85.39 (C2), 101.68 (C1), 109.51, 113.50 (Cq iso-prop), 128.60 129.65 129.95, 133.62 (Carom *O*-Bz), 165.06 (C=O, *O*-Bz); HRMS(ESI) *m/z* calcld. for $[C_{19}H_{24}NaO_7]^+$: 387.1414, obsd.: 387.1413.

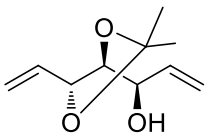


Benzoyl 2,3-*O*-isopropylidene-α-D-lyxo-hex-5-enofuranose (5.22). Benzoyl **5.25** was dissolved in glacial acetic acid (32 mL), water (14 mL) was added and the resulting mixture was stirred at 55°C. After 75 min. TLC analysis (PE/EtOAc, 2/1, v/v) showed complete disappearance of starting material and the formation of a lower running product (R_f 0.53) and concentrated *in vacuo*. The crude was suspended in triethyl ortho formate (25 mL) and glacial acetic acid (0.4 mL, 0.7 mmol) was added and left until full conversion was observed by t.l.c analysis (R_f 0.76, PE/EtOAc, 2/1, v/v). The solution was concentrated and heated to 170 °C in an open flask. Diphenyl acetic acid (212 mg, 0.06 mmol, 0.44 mol%) was added and after 4 hours all starting material had been transformed into a higher running product (R_f 0.55, PE/EtOAc, 4/1, v/v). The resulting brown oil was allowed to cool to room temperature. Purification by flash chromatography (silica gel, PE/EtOAc 7.5/1) gave homogenous title compound **5.22** in a 80% yield (3.2 g, 11.0 mmol). R_f 0.53 (PE/EtOAc, 2/1, v/v): ^1H NMR (400 MHz, $CDCl_3$); δ 1.36, 1.53 (2x s, 3H, Me iso-prop), 4.62-4.66 (m, 1H), 4.82-4.86 (m, 1H), 4.87 - 4.90 (m, 1H), 5.36 (d, *J*=

10.3 Hz, 1H, H6-*cis*), 5.44 (d, *J*= 17.6 Hz, 1H, H6-*trans*), 6.01 (m, *J*=17.6, 10.3 Hz, 1H, H5), 6.42 (s, 1H, H1), 7.40 - 7.49 (m, 2H, H_{arom} *O*-Bz), 7.55 - 7.62 (m, 1H, H_{arom} *O*-Bz), 8.00 - 8.04 (m, 2H, H_{arom} *O*-Bz); ^{13}C NMR (101 MHz, $CDCl_3$); δ 25.22, 26.25 (2x Me iso-prop), 81.72, 81.80, 85.63 (C2, C3, C4), 101.72 (C1), 119.29 (C6), 128.64, 129,39, 130.36 (Carom *O*-Bz), 132.16 (C5), 133.92 (Carom *O*-Bz). HRMS(ESI) *m*/*z* calcld. for $[C_{16}H_{18}NaO_5]^+$: 313.1046, obsd.: 313.1040.

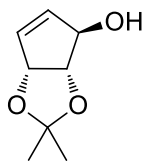


2,3-*O*-isopropylidene-α-D-lyxo-hex-5-enofuranose (5.26). To a solution of alkene **5.22** in MeOH, *t*-BuOK was added until a pH of 12-14 and left to stir at room temperature. after 4.5 hour TLC analysis (PE/EtOAc, 1/1, v/v) showed complete disappearance of the starting material and formation of a lower running product (R_f 0.66). DOWEX H^+ was added until neutral (pH 7) by pH paper and concentrated. Purification by flash chromatography (silica gel, PE/EtOAc 10/1→7.5/1) gave homogenous title compound **5.26** in a 95% yield (3.2 g, 11.0 mmol). R_f 0.66 (PE/EtOAc, 1/1, v/v): IR (v): 494.6, 515.2, 556.2, 717.9, 768.7, 823.0, 850.3, 877.4, 905.4, 842.3, 992.8, 1014.8, 1039.5, 1069.0, 1161.7, 1210.1, 1237.46, 1377.8, 2943.0, 3468.3 cm^{-1}; ^{1}H NMR (400 MHz, $CDCl_3$); δ 1.32, 1.47 (2x s, 3H, Me iso-prop), 4.59 - 4.63 (m, 1H, H4), 4.64 (d, *J*=5.9 Hz, 1H, H2), 4.73 (dd, *J*= 3.5 5.9 Hz, 1H, H3), 5.34 (d, *J*= 10.2 Hz, 1H, H6-*cis*), 5.41 (s, 1H, H1), 5.42 (d, *J*= 17.2 Hz, 1H, H6-*trans*), 5.99 (ddd, *J*= 7.1, 10.5, 17.2 Hz, 1H, H5); ^{13}C NMR (101 MHz, $CDCl_3$); δ 25.03, 26.22 (2x Me iso-prop), 81.67, 81.70 (C3, C4), 85.95 (C2), 101.25 (C1), 112.80 (Cq iso-prop), 119.36 (C6), 132.34 (C5). HRMS(ESI) *m*/*z* calcld. for $[C_9H_{14}NaO_4]^+$: 209.0784, obsd.: 209.0788.

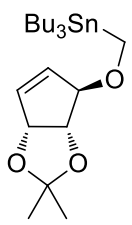


(3*R*,4*S*,5*R*)-3,4-*O*-isopropylidenehepta-1,6-diene-3,4,5-triol (5.21). To methyl triphenylphosphonium bromide (11.6 g, 32.5 mmol) in THF/HMPA (40 mL, 3/1, v/v), KOtBu (3.6 g, 32.5 mmol) was added at -15°C. The resulting solution was stirred for an additional hour at 0°C. A solution of lactol **5.26** (1.72 mg, 9.34 mmol) in THF (2 mL) was added and the solution was allowed to warm to room temperature. After 14 hours all starting material had been converted into a slightly higher running product as indicated by t.l.c analysis (R_f 0.69, PE/EtOAc, 1/1, v/v). The reaction mixture was quenched with aq. sat. NH_4Cl and extracted with Et_2O. The ether layer was

dried (MgSO4), filtered and concentrated. Purification by flash chromatography (silica gel, PE/EtOAc 10/1→8/1) gave dialkene **5.21**. R_f 0.69 (PE/EtOAc, 1/1, v/v): IR (v): 513.2, 721.9, 871.4, 928.4, 1043.5, 1216.7, 1243.1, 3475.7 cm^{-1}): ^1H NMR (400 MHz, $CDCl_3$); δ 1.40, 1.54 (2x s, 3H, Me iso-prop), 4.06-4.16 (m, 2H), 4.48-4.64 (m, 1H), 5.23 (d, *J*= 10.6 Hz, 1H), 5.30 (d, *J*= 10.2 Hz, 1H), 5.37 (d, *J*= 17.2 Hz, 1H), 5.38 (d, *J*= 17.2 Hz, 1H), 5.85 (ddd, *J*= 5.1, 10.6, 17.2 Hz, 1H), 6.02 (ddd, *J*= 8.2, 10.2, 17.2 Hz, 1H); ^{13}C NMR (101 MHz, $CDCl_3$); δ 25.16, 27.56 (2x Me iso-prop), 70.77, 79.12, 80.82 (C3, C4, C5), 108.97 (Cq iso-prop), 117.19, 119.55 (C1, C6), 134.11, 136.93 (C2, C5). HRMS(ESI) *m/z* calcld. for $[C_{10}H_{16}NaO_3]^+$: 207.0992, obsd.: 207.0986.

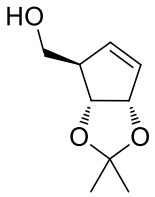


(3*R*,4*S*,5*R*)-3,4-*O*-Isopropylidene-cyclopentene-3,4,5-triol (5.20). Dialkene **5.21** was dissolved in CH_2Cl_2 (25 mL), the solution was degassed by passing through a stream of argon for 15 min and 36 mg (0.5 mol%) of Grubbs' 1st generation catalyst was added under an atmosphere of argon. After 5 hours TLC analysis revealed complete conversion (R_f 0.47, PE/EtOAc, 1/1, v/v). The reaction mixture was purified by flash chromatography (silica gel, PE/EtOAc 6/1→2/1) then gave cyclopentene **5.20** in 52% yield over two steps (760 mg, 4.87 mmol) as a brown oil. R_f 0.47, (PE/EtOAc, 1/1, v/v) IR (v): 722.0, 865.5, 1039.4, 1159.5, 1209.8, 2933.9, 3390.2 cm^{-1}): ^1H NMR (400 MHz, $CDCl_3$); δ 1.34, 1.39 (2x s, 3H, Me iso-prop), 1.93 (bs., 1H, OH), 4.52 (d, *J*= 5.6 Hz, 1H), 4.79 (bs, 1H), 5.28 (d, *J*= 5.6 Hz, 1H), 5.88 - 5.93 (m, 1H), 6.02-6.04 (m, 1H); ^{13}C NMR (101 MHz, $CDCl_3$); δ 25.87, 27.48 (Me iso-prop), 81.16, 84.42, 86.08, 111.88 (Cq iso-prop), 134.77, 135.77.



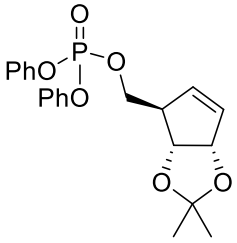
(3*R*,4*S*,5*R*)-3,4-O-Isopropylidene-5-methyltributylstannyl- cyclopentene-3,4,5-triol (5.27). Alcohol **5.20** (317.4 mg, 2.03 mmol) was dissolved in THF (7 mL) and cooled to 0 °C. Potassium hydride (165 mg, 4 mmol) was added and the resulting reaction mixture left for 10 min. Bu3-SnCH2I (0.66 mL, 2.2 mmol) was added and the solution was allowed to warm to rt. Stirring was continued for 60 min after which time TLC analysis (PE/EtOAc, 9/1, v/v) indicated consumption of all starting material and the formation of a higher running product (R_f 0.8). Water was added diluted

with EtOAc, the organic phase was washed with H_20 (2x), brine (1x), dried over $MgSO_4$, filtered and concentrated. Purification by flash chromatography (silica gel, 4% EtOAc in PE) gave stannane ether **5.27** as a pale yellow oil. R_f 0.80 (PE/EtOAc, 9/1, v/v). ^{1}H NMR (400 MHz, $CDCl_3$) δ: 0.86-0.91 (m, 9H, Me $SnBu_3$), 1.22–1.34 (m, 6H, CH_2 $SnBu_3$), 1.36, 1.41 (2x s, Me iso-prop), 1.46-1.52 (m, 6H, CH_2 $SnBu_3$ 3.69), 3.69 (d, *J*= 10.0 Hz, 1H, OCH_2Sn), 3.88 (d, *J*=10.0 Hz, 1H, OCH_2 Sn), 4.24 (s, 1H), 4.51 (d, *J*= 6.0 Hz, 1H), 5.19- 5.22 (m 1H), 5.88-5.91 (m, 1H), 5.97-6.00 (m, 1H); ^{13}C NMR (101 MHz, $CDCl_3$); δ 9.17, 13.86 ($SnBu_3$), 25.83, 27.48 (2x Me iso-prop), 27.42, 29.26 ($SnBu_3$), 60.49 (OCH_2Sn), 83.06, 84.43, 92.91 (C-3, C-4, C-5), 111.71 (Cq iso-prop), 133.72, 134.11 (C-1, C-2).



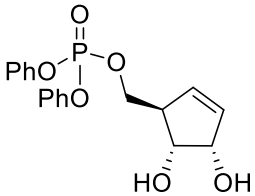
(3*S*,4*R*,5*R*)-3,4-O-Isopropylidene-5-hydroxymethylcyclo-pentene-3,4-diol (5.19)

Stannane ether **5.27** (473.3 mg, 1.03 mmol) was dissolved in THF (20 mL) and cooled to -78 ^{o}C. *n*-BuLi (1.1 mL, 1.42 M in THF, 1.6 mmol) was added dropwise via the cold wall of the flask. The solution was allowed to warm to rt. After 2 hours TLC analysis (PE/EtOAc, 2/1, v/v) showed formation of a lower running product (R_f 0.2). The solution was poured onto sat. aq. NH_4Cl and extracted with EtOAc (2x). The combined organic layers were washed with brine (1x) dried over $MgSO_4$, filtered and concentrated. Purification by flash chromatography (silica gel, PE/EtOAc 1/1) gave cyclopentene **5.19** as a clear oil in 95% yield (117 mg, 1.0 mmol). R_f 0.2 (PE/EtOAc, 2/1, v/v). IR (v): 512.2, 738.8, 859.8, 938.9, 1042.2, 1158.1, 1206.9, 1370.7, 2930.9, 3405.1 cm^{-1}: ^{1}H NMR (400 MHz, CDCl3): δ 1.36, 1.42 (2x s, 3H, Me iso-prop), 2.95-3.02 (m, 1H, H5), 3.56-3.61 (m, 1H, CH_2OH), 3.72-3.78 (m, 1H, CH_2OH), 4.59 (d, *J*= 6.0 Hz, 1H, H3), 5.12-5.17 (m, 1H, H-4), 5.73-5.77 (m, 1H, H2), 5.91-5.96 (m, 1H, H1); ^{13}C-NMR (100 MHz, $CDCl_3$): δ 25.64, 27.46 (2x Me iso-prop), 54.81 (C5), 63.72 (CH_2OH), 81.49 (C4), 85.16 (C3), 110.09 (Cq iso-prop), 132.9 (C2), 133.5 (C1). HRMS(ESI) *m*/*z* calcld. for $[C_9H_{14}NaO_3]^+$: 193.0835, obsd.: 193.840

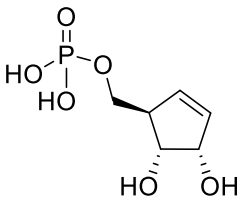


(3*S*,4*R*,5*R*)-3,4-O-Isopropylidene-5-(hydroxymethyl-*O*-phosphoryl)-pentene-3,4-diol (5.28). Hydroxymethyl **5.19** (30 mg, 0.176 mmol) was dissolved in CH_2Cl_2 (1 mL) and cooled to 0 ^{o}C. triethyl amine (73 µL, 0.53 mmol) was added followed by diphenyl phosphoryl chloride (43 µL, 0.21 mmol) and allowed to reach rt, after 4 hours complete consumption of

starting material into a higher running product was observed by TLC analysis (PE/EtOAc, 1/1, v/v, R_f 0.60). The reaction was quenched with water. The organic layer was washed with water (1x), sat. aq. $NaHCO_3$ (1x), dried over $MgSO_4$, filtered and concentrated. Purification by flash chromatography (silica gel, PE/EtOAc 4/1) gave protected phosphate **5.19** as a clear oil in 84% yield (59 mg, 1.0 mmol). R_f 0.2 (PE/EtOAc, 2/1, v/v). IR (v): 505.4, 688.4, 754.5, 945.9, 1021.1, 1160.7, 1187.7, 1284.6, 1370.4, 1487.8, 1589.8, 1723.8 cm^{-1}: ^1H NMR (400 MHz, CDCl3): δ 1.29, 1.38 (2x s, 3H, Me iso-prop), 3.09-3.13 (m, 1H, H5), 4.13-4.17 (m, 1H, CH_2OH), 4.27-4.32 (m, 1H, CH_2OH), 4.47 (d, *J*=6.0 Hz, 1H, H4), 4.97-4.99 (m, 1H, H-4), 5.65-5.67 (m, 1H, H2), 5.83-5.85 (m, 1H, H1); ^{13}C-NMR (100 MHz, $CDCl_3$): δ 25.49, 27.26 (2x Me iso-prop), 52.63 (C5), 69.25 (CH_2OH), 80.52 (C4), 84.85 (C3), 110.26 (Cq iso-prop), 119.99, 120.00, 120.03, 120.06, 125.45, 128.77, 129.78, 129.80, 130.85 (Carom Bn), 131.64 (C1), 133.95 (C2),132.44, 150.40, 150.49 (Carom Bn); ^{31}P NMR (162 MHz, $CDCl_3$); δ 12.09 (s, 1P); HRMS(ESI) *m/z* calcld. for $[C_{21}H_{24}O_6P]^+$: 403.1305, obsd.: 403.1335



(3*S*,4*R*,5*R*)-5-(hydroxymethyl-*O*-diphenylphosphoryl)-cyclo-pentene-3,4-diol (5.29). Protected phosphate **5.28** (31.5 mg, 0.078 mmol) was taken op in methanol (2 mL) and acidified with concentrated HCl (0.2 mL). The resulting solution was left for 1 min and then concentrated to dryness by rotary evaporation. TLC analysis of the reaction showed complete consumption of starting material and a major product (CH_2Cl_2/MeOH, 95/5, v/v, R_f 0.5). Purification by flash chromatography (silica gel, CH_2Cl_2/MeOH 100/0 → 90/10) gave 2,3-hydroxy **5.29** as a clear oil in 70% yield (20 mg, 0.05 mmol). (CH_2Cl_2/MeOH, 95/5, v/v, R_f 0.5) IR (v): 505.4, 688.4, 754.5, 945.9, 1021.1, 1160.7, 1187.7, 1284.6, 1370.4, 1487.8, 1589.8, 1723.8, 2930.4 cm^{-1}: ^1H NMR (400 MHz, CDCl3): δ 2.72 (bs, 2H, OH-2, OH-3), 2.96-2.97 (m, 1H, H5), 3.97-4.00 (m, 1H, H4), 4.24-4.35 (m, 2H, CH_2OH), 4.48-4.49 (m, 1H, H4), 5.79-5.81 (m, 1H, H1), 5.85-5.88 (m, 1H, H2); ^{13}C-NMR (100 MHz, $CDCl_3$): δ 52.34 (d, *J*=6.8 Hz, C5), 69.22(d, *J*=6.1 Hz, CH_2OH), 72.85 (C4), 74.68 (C3), 120.00, 120.06, 125.52, 129.84 (C Ph), 133.61, 133.70 (C1, C2), 150.34, 150.42 (C_{arom} Ph); ^{31}P NMR (162 MHz, $CDCl_3$); δ 12.00 (s, 1P);. HRMS(ESI) *m/z* calcld. for $[C_{18}H_{20}O_6P]^+$: 363.0992, obsd.:363.1032



(3*S*,4*R*,5*R*)-5-(hydroxymethyl-*O*-phosphoryl)-cyclo-pentene-3,4-diol (5.3). Diol **5.29** (13.2 mg, 0.036 mmol) was taken op in THF (2 mL) and cooled down to -78 °C. A flow of ammonium gas was blown over the flask until approximately 10 mL was collected. Sodium metal cut clean of oxydised surface (5-10 mg cubes) were added until a dark blue colour persisted. This was maintained for 45 min after which MeOH was added (3 mL) and the ammonia blown off under a flow of nitrogen. The reaction was then neutralised with DOWEX® resin, the resin filtered off and was washed with methanol, and methanol was removed on a rotary evaporator. Purification by reverse phase chromatography (C18, H_2O) gave cyclopentene phosphate **5.3** as a clear oil in 55% yield (4.2 mg, 0.02 mmol). IR (ν): 795.0, 960.8, 1087.1, 1333.6, 1428.9, 1675.7, 3157.5 cm^{-1}: ^1H NMR (400 MHz, D_2O): δ 2.77-2.78 (m, 1H, H5), 3.79-3.94 (m, 1H, CH_2OH), 4.44 (d, *J*= 5.2 Hz, 1H, H4), 5.78-5.85 (m, 1H, H-3), 7.05-7.08 (m, 1H), 7.24-7.28 (m, 1H)(H1, H2). HRMS(ESI) *m/z* calcld. for $[C_6H_{12}O_6P]^+$: 211.0366, obsd.:211.0370

8.5.3 Inhibition study of transition state analogues

The cyclopentene phosphate, iminoribitol phosphate and iminoribitol phosphonate were tested for inhibition again the ATP-PRTase enzyme from *M. tuberculosis* and *C. jejuni*. The compounds were tested under competitive conditions ranging from 0 to 500 µM of transition state analogue. The inhibition was tested under varying PRPP concentrations (1 to 3 times K_m, 50-150 µM for the *M. tuberculosis* ATP-PRTase enzyme and 11-33 µM for the *C. jejuni* ATP-PRTase enzyme) the ATP concentrations was kept constant at 3 times *K*m (750 µM *M. tuberculosis* ATP-PRTase and 500 µM *C. jejuni* ATP-PRTase enzyme). The reactions were preincubated for at least two minutes prior to initiation with PRPP.

Data obtained from the kinetic assays where processed using Prism 6.0 and the standard calculations for a competitive inhibitor were used to find inhibition constants.

Concentration of the phosphate containing analogues were established by cleaving the phosphate with alkaline phosphatase followed by assessment of the released inorganic phosphate by lenzetta assay (section 8.5.1). The concentration for the iminoribitol phosphonate was found by ^{31}P NMR with phosphoric acid as the internal standard.

8.6 Experimental procedures for chapter 6

8.6.1 Purification of C. jejuni core mutant

Expression and purification of *C. jejuni* ATP-PRTase core mutant was carried out according to established procedures by Dr G. H. mittlestäd.[67] The constructuct contained a GST tag. The H33 and H35 allowed partial purification by HiTrap column prior to further purification by GST column chromatography. *C. jejuni* ATP-PRTase core mutant/pDEST15/*E. coli* Bl21* (DE3) cell cultures were grown in autoclaved LB (Lennox LB) at 20 g/L containing the appropriate concentration of Ampicillin. Pre-cultures in LB (Lennox LB) media containing Ampicillin (Section 8.1.21) were inoculated using a single culture grown from a glycerol stock and grown overnight with shaking (180 rpm) at 37 °C. Approximately 25 mL of pre-culture per litre of cell culture were used to inoculate growths (OD_{600} of 0.1 AU). Cells were grown at 37 °C until mid-logarithmic phase (OD_{600} of 0.4–0.6 AU). The protein expression was induced by addition of isopropyl β-D-1-thiogalactopyranoside (IPTG, 0.5 mM final concentration), and growth was continued for 4 hours. Cells were harvested by centrifugation (14,000 *g* for 15 min at 4 °C) and the cell pellet was flash-frozen in liquid nitrogen and stored at −80 °C. The cell pellet was resuspended in low imidazole buffer (50 mM Na_2HPO_4, 500 mM NaCl, 5 mM $MgCl_2$ and 20 mM imidazole at pH 8.5), lysed for 3x 5min 80 %, spun down for 30 min at 40,000 *g*. The supernatant was filtered and loaded onto a 5 mL GE Healthcare His trap HP column. Protein which bound to the column was eluted with high imidazole buffer (50 mM Na_2HPO_4, 500 mM NaCl, 5 mM $MgCl_2$ and 500 mM imidazole at pH 8.5) and fractions containing enzyme determined by SDS-page gel electrophoresis were pooled together. The pooled fractions were buffer exchanged by spin concentrator into phosphate buffered seline (PBS) at pH 7.3. The target protein was then loaded onto a 5 mL GE Healthcare GST HP column and eluted with 10 mM glutathione in a 50 mM Tris buffer at pH 8.0. Fractions containing the *C. jejuni* ATP-PRTase core mutant enzyme were pooled together and buffer exchanged back into PBS buffer by spin concentrator and treated with TEV protease at a 1% mol/mol concentration at 37 °C for one hour and then at at 4°C overnight. TEV treated enzyme was run through a 5 mL GE Healthcare GST HP column and the flow through was collected, concentrated and buffer exchanged into Tris buffer (100mM NaCl, 5mM $MgCl_2$, and 50 mM Tris at pH 8.0 before loading onto a 26/60 size exclusion column preequilibrated with Tris buffer. Fractions containing *C. jejuni* ATP-PRTase core mutant as detected by sds-

page gel electrophoresis were concentrated to 15-20 mg/mL, flash-frozen in liquid N_2 and stored at -80 °C. 4 litres of culture would yield approximately 20-30 mg of protein per purification.

8.6.2 Protein stability assay

Enzyme stability was monitored by the rate of conversion of PRPP and ATP into PRATP and PP_i. Because NMR requires the longest experimental time, conditions similar to these samples were chosen for the stability assay (Protein concentration 15 µM, Tris 10 mM, NaCl 30 mM, $MgCl_2$ 1 mM and D_2O 10% at pH 8.5). The sample was kept at room temperature for the duration of the activity assay. The rate of conversion was measured after the sample was made up (day 0) and monitored with 24 hours interval for 6 days, a reduced conversion rate compared to this initial value (day 0) was considered a sign of degradation of the enzyme (Figure 8.3). On day one no reduction in the conversion rate was observed, after day two a 7% loss in conversion was observed and over the course of the next 4 days this progressed with a total loss of 1/3 of the activity over that time. These results show that stability of the protein will not be a problem for the required experiment time.

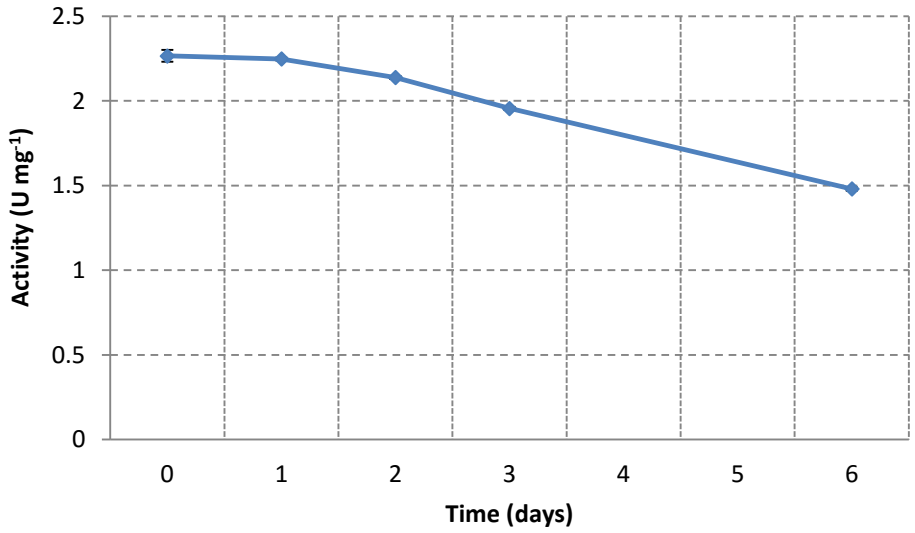


Figure 8.3. The graph displays the rate of conversion in (U mg^{-1}) of PRPP (0.3 mM) and ATP (2.5 mM) into PR-ATP and PP_i over a 6 day time period. Reduction in the conversion rate was considered a sign of degradation.

8.6.3 STD-NMR sample preparation

The STD-NMR spectra were acquired by Dr Herbert Wong from Callaghan Inovation on a DRX 500 spectrometer (Bruker) operating at 500MHz and 25 °C. Samples consisted of 5 mM $MgCl_2$, 100 mM NaCl, 15% (v/v) D_2O and 10 mM Tris (HCl) at pH 8.0. Ligands were added to a concentration of 2mM from 100mM stock made up in D_2O or DMSO-D6 depending on solubility. DPFGSE water suppression was applied to all experiments. A ^{1}H-Proton and STD-NMR spectrum was taken in absence of enzyme. *C. jejuni* wt ATP-PRTase Enzyme was then added to a finial concentration of 15 µM. A second STD-NMR spectrum was then taken and compared to reference spectra. The STD pulse program followed the standard provided scheme with the topspin software package and with optimisation of the pre-acquisition time from 5s to 3s, a protein saturation time of 2 seconds and the on-resonance set to 11 or -1ppm. Spectra were collected with 200-512 scans.

8.6.4 Differential scanning fluorimetry

Differential scanning fluorimetry was carried out on an iCycler iQ5 Multicolour Real-Time PCR Detection System (Bio-Rad). The melting temperature of protein in presence of natural and non-natural ligands was established following a modified method by Ericsson *et al.*[210] A typical run was set up measuring the temperature from 20 °C to 80 °C at 0.2 °C increments with a 20s dwell time. SYPRO orange dye 250X (Invitrogen) was used to visualise the melting profile.

Samples were measured in three replicates and one blank. Samples and dye were prepared on ice. A typical sample was set up by the addition of 5 µL of 250 times SYPRO orange dye to 100 µL of buffer (100 mM NaCl, 5mM $MgCl_2$ and 50 mM Tris at pH 8.0), which included the appropriate ligand or other additive. Of this stock 21 µL was added to the blank well and 4 µL of Buffer was added. To the remaining sample stock 16 µL of protein (0.5 – 1 mg/mL) was added, and 25 µL dispensed into the three replicate wells

For the analysis of the data an Excel spreadsheet with custom VBA-scripted macro created by Dr T. Allison was used to display melting profiles. The melting temperature were determent from the maximum inflection of the melting curve, corrected with the blank analysis.

Referencing

1. Garrett, R. H. & Grisham, C. M. in *Biochemistry* 405–511 (Thomson Brooks/cole, 2005).
2. Schramm, V. L. Enzymatic transition states and transition state analog design. *Annu. Rev. Biochem.* **67,** 693–720 (1998).
3. Ioerger, T. R. & Sacchettini, J. C. Structural genomics approach to drug discovery for Mycobacterium tuberculosis. *Curr. Opin. Microbiol.* **12,** 318–325 (2009).
4. Lamichhane, G. Novel targets in *M. tuberculosis*: search for new drugs. *Trends Mol. Med.* **17,** 25–33 (2010).
5. World Health Organization. *Global Tuberculosis Report 2015*. *WHO* (2015). doi:10.1007/s13398-014-0173-7.2
6. Cambau, E. & Drancourt, M. Steps towards the discovery of Mycobacterium tuberculosis by Robert Koch, 1882. *Clin. Microbiol. Infect.* **20,** 196–201 (2014).
7. Jawahar, M. S. *et al.* Treatment of lymph node tuberculosis--a randomized clinical trial of two 6-month regimens. *Trop. Med. Int. Heal. TM IH* **10,** 1090–1098 (2005).
8. Turgut, M. Spinal tuberculosis (Pott's disease): its clinical presentation, surgical management, and outcome. A survey study on 694 patients. *Neurosurg. Rev.* **24,** 8–13 (2001).
9. Silva Miranda, M., Breiman, A., Allain, S., Deknuydt, F. & Altare, F. The Tuberculous Granuloma: An Unsuccessful Host Defence Mechanism Providing a Safety Shelter for the Bacteria? *Clin. Dev. Immunol.* **2012,** 1–14 (2012).
10. Russell, D. G. Who puts the tubercle in tuberculosis? *Nat. Rev. Microbiol.* **5,** 39–47 (2007).
11. Flynn, J. L. & Chan, J. Tuberculosis : Latency and Reactivation. *Infect. Immun.* **69,** 4195–4201 (2001).
12. Koul, A., Arnoult, E., Lounis, N., Guillemont, J. & Andries, K. The challenge of new drug discovery for tuberculosis. *Nature* **469,** 483–490 (2011).
13. Lönnroth, K. *et al.* Tuberculosis control and elimination 2010–50: cure, care, and social development. *Lancet* **375,** 1814–1829 (2010).
14. World Health Organization. Tuberculosis. *Fact sheet N 104* at <http://www.who.int/mediacentre/factsheets/fs104/en/>
15. Gandhi, N. R. N. N. R. *et al.* Multidrug-resistant and extensively drug-resistant tuberculosis: a threat to global control of tuberculosis. *Lancet* **375,** 1830–1843 (2010).
16. World Health organization. *Global Tuberculosis Report 2014, WHO*. (2014). doi:10.1007/s13398-014-0173-7.2
17. Uplekar, M. *et al.* WHO's new End TB Strategy. *Lancet* **385,** 1799–1801 (2015).
18. Dalton, T. *et al.* Prevalence of and risk factors for resistance to second-line drugs in people with multidrug-resistant tuberculosis in eight countries: a prospective cohort study. *Lancet* **380,** 1406–1417 (2012).
19. van den Boogaard, J., Kibiki, G. S., Kisanga, E. R., Boeree, M. J. & Aarnoutse, R. E. New Drugs against Tuberculosis: Problems, Progress, and Evaluation of Agents in Clinical Development. *Antimicrob. Agents Chemother.* **53,** 849–862 (2009).
20. Zumla, A., Nahid, P. & Cole, S. T. Advances in the development of new tuberculosis drugs and treatment regimens. *Nat. Rev. Drug Discov.* **12,** 388–404 (2013).
21. Upton, A. M. *et al.* In Vitro and In Vivo Activities of the Nitroimidazole TBA-354 against *Mycobacterium tuberculosis*. *Antimicrob. Agents Chemother.* **59,** 136–144 (2015).
22. Skripconoka, V. *et al.* Delamanid improves outcomes and reduces mortality in multidrug-resistant tuberculosis. *Eur. Respir. J.* **41,** 1393–1400 (2013).
23. Matsumoto, M. *et al.* OPC-67683, a nitro-dihydro-imidazooxazole derivative with promising action against tuberculosis in vitro and in mice. *PLoS Med.* **3,** 2131–2144 (2006).
24. Matteelli, A., Carvalho, A. C., Dooley, K. E. & Kritski, A. TMC207: the first compound of a new class of potent anti-tuberculosis drugs. *Futur. Microbiol.* **5,** 849–858 (2010).
25. Ge, B., Wang, F., Sjölund-Karlsson, M. & McDermott, P. F. Antimicrobial resistance in Campylobacter: Susceptibility testing methods and resistance trends. *J. Microbiol. Methods*

95, 57–67 (2013).
26. Rees, J. H., Soudain, S. E., Gregson, N. A. & Hughes, R. A. . *Campylobacter jejuni* Infection and Guillain-Barré Syndrome. *N. Engl. J. Med.* **333,** 1374–1379 (1995).
27. Altekruse, S. F., Stern, N. J., Fields, P. I. & Swerdlow, D. L. *Campylobacter jejuni* - An Emerging Foodborne Pathagen. *Emerg. Infect. Dis.* **5,** 28–36 (1999).
28. Allos, B. M. *Campylobacter jejuni* Infections: Update on Emerging Issues and Trends. *Food Saf.* 1201–1206 (2001).
29. Hooper, D. C. Emerging Mechanisms of Fluoroquinolone Resistance. *Emerg. Infect. Dis.* **7,** 337–341 (2001).
30. Jastrzebowska, K. & Gabriel, I. Inhibitors of amino acids biosynthesis as antifungal agents. *Amino Acids* **47,** 227–249 (2015).
31. Abdo, M.-R. *et al.* Anti-virulence strategy against Brucella suis: synthesis, biological evaluation and molecular modeling of selective histidinol dehydrogenase inhibitors. *Org. Biomol. Chem.* **9,** 3681–3690 (2011).
32. Parish, T. Starvation survival response of Mycobacterium tuberculosis. *J Bacteriol* **185,** 6702–6706 (2003).
33. Bange, F. C., Brown, a M. & Jacobs, W. R. Leucine auxotrophy restricts growth of *Mycobacterium bovis* BCG in macrophages. *Infect. Immun.* **64,** 1794–1799 (1996).
34. Hondalus, M. K. *et al.* Attenuation of and protection induced by a leucine auxotroph of Mycobacterium tuberculosis. *Infect. Immun.* **68,** 2888–2898 (2000).
35. Fani, R., Mori, E. & Lazcano, A. Histidine Biosynthetic Pathway and Genes: Structure, Regulation, and Evolution. *Orig. Life Evol. Biosph.* **26,** 1–26 (1996).
36. Kulis-Horn, R. K., Persicke, M. & Kalinowski, J. Histidine biosynthesis, its regulation and biotechnological application in *Corynebacterium glutamicum*. *Microb. Biotechnol.* **7,** 5–25 (2014).
37. Martin, G. The First Enzyme in Histidine Biosynthesis : The Nature Feedback Inhibition by Histidine. *J. Biol. Chem.* **238,** 257–268 (1963).
38. Sinha, S. C. & Smith, J. L. The PRT protein family. *Curr. Opin. Struct. Biol.* **11,** 733–739 (2001).
39. Lee, C. E., Goodfellow, C., Javid-Majd, F., Baker, E. N. & Shaun Lott, J. The crystal structure of TrpD, a metabolic enzyme essential for lung colonization by *Mycobacterium tuberculosis*, in complex with its substrate phosphoribosylpyrophosphate. *J. Mol. Biol.* **355,** 784–797 (2006).
40. Schumacher, M. a *et al.* Crystal structures of *Toxoplasma gondii* uracil phosphoribosyltransferase reveal the atomic basis of pyrimidine discrimination and prodrug binding. *EMBO J.* **17,** 3219–3232 (1998).
41. Phillips, C. L., Ullman, B., Brennan, R. G. & Hill, C. P. Crystal structures of adenine phosphoribosyltransferase from Leishmania donovani. *EMBO J.* **18,** 3533–3545 (1999).
42. Muchmore, C. R., Krahn, J. M., Kim, J. H., Zalkin, H. & Smith, J. L. Crystal structure of glutamine phosphoribosylpyrophosphate amidotransferase from *Escherichia coli*. *Protein Sci.* **7,** 39–51 (1998).
43. Malik, S. S., Patterson, D. N., Ncube, Z. & Toth, E. a. The crystal structure of human quinolinic acid phosphoribosyltransferase in complex with its inhibitor phthalic acid. *Proteins* **82,** 405–414 (2014).
44. Sharma, V., Grubmeyer, C. & Sacchettini, J. C. Crystal structure of quinolinic acid phosphoribosyltransferase from *Mycobacterium tuberculosis*: a potential TB drug target. *Structure* **6,** 1587–1599 (1998).
45. Eads, J. C., Ozturk, D., Wexler, T. B., Grubmeyer, C. & Sacchettini, J. C. A new function for a common fold: the crystal structure of quinolinic acid phosphoribosyltransferase. *Structure* **5,** 47–58 (1997).
46. Cookson, T. V. M. *et al.* Structures of *Mycobacterium tuberculosis* Anthranilate Phosphoribosyltransferase Variants Reveal the Conformational Changes That Facilitate Delivery of the Substrate to the Active Site. *Biochemistry* **54,** 6082–6092 (2015).

47. Cho, Y., Sharma, V. & Sacchettini, J. C. Crystal structure of ATP phosphoribosyltransferase from *Mycobacterium tuberculosis*. *J. Biol. Chem.* **278,** 8333–8339 (2003).
48. Sissler, M. *et al.* An aminoacyl-tRNA synthetase paralog with a catalytic role in histidine biosynthesis. *Proc. Natl. Acad. Sci. U.S.A.* **96,** 8985–8990 (1999).
49. Lohkamp, B., McDermott, G., Campbell, S. a., Coggins, J. R. & Lapthorn, A. J. The Structure of *Escherichia coli* ATP-phosphoribosyltransferase: Identification of Substrate Binding Sites and Mode of AMP Inhibition. *J. Mol. Biol.* **336,** 131–144 (2004).
50. Lang, E. J. M., Cross, P. J., Mittelsteadt, G., Jameson, G. B. & Parker, E. J. Allosteric ACTion : the varied ACT domains regulating enzymes of amino-acid metabolism. *Curr. Opin. Struct. Biol.* 102–111 (2014).
51. Kryvi, H. & Klungsöyr, L. Kinetic properties of phosphoribosyladenosine triphosphate synthetase inhibition by aggregation at high enzyme concentrations. *Biochim. Biophys. Acta* **235,** 429–434 (1971).
52. Morton, D. P. & Parsons, S. M. Inhibition of ATP phosphoribosyltransferase by AMP and ADP in the absence and presence of histidine. *Arch. Biochem. Biophys.* **181,** 643–648 (1977).
53. Champagne, K. S., Sissler, M., Larrabee, Y., Doublié, S. & Francklyn, C. S. Activation of the hetero-octameric ATP phosphoribosyl transferase through subunit interface rearrangement by a tRNA synthetase paralog. *J. Biol. Chem.* **280,** 34096–104 (2005).
54. Vega, M. C. *et al.* Regulation of the hetero-octameric ATP phosphoribosyl transferase complex from *Thermotoga maritima* by a tRNA synthase-like subunit. *Mol. Microbiol.* **55,** 675–686 (2005).
55. Champagne, K. S., Piscitelli, E. & Francklyn, C. S. Substrate recognition by the hetero-octameric ATP phosphoribosyltransferase from *Lactococcus lactis*. *Biochemistry* **128,** 14691–14696 (2006).
56. Morton, D. P. & Parsons, S. M. Biosynthetic direction substrate kinetics and product inhibition studies on the first enzyme of histidine biosynthesis, adenosine triphosphate phosphoribosyltransferase. *Arch. Biochem. Biophys.* **175,** 677–686 (1976).
57. Hughson, F. M., Schedl, P. & Smith, J. L. Forming and inhibiting PRT active sites. *Nat. Struc. Biol.* **6,** 502–504 (1999).
58. Parsons, S. M. P., Chelsky, D. & Goitein, R. K. Primary 14C and a Secondary 3H Substrate Effects for Some Phosphoribosyltransferases. *J. Biol. Chem.* **253,** 2963–2971 (1978).
59. Cleland, W. W. The use of isotope effects to determine enzyme mechanisms. *Arch. Biochem. Biophys.* **433,** 2–12 (2005).
60. Tao, W., Grubmeyer, C. & Blanchard, J. S. Transition state structure of *Salmonella typhimurium* orotate phosphoribosyltransferase. *Biochemistry* **35,** 14–21 (1996).
61. Zhang, Y., Luo, M. & Schramm, V. L. Transition states of *Plasmodium falciparum* and human orotate phosphoribosyltransferases. *J. Am. Chem. Soc.* **131,** 4685–4694 (2009).
62. Zhang, Y. & Schramm, V. L. Pyrophosphate interactions at the transition states of *Plasmodium falciparum* and human orotate phosphoribosyltransferases. *J. Am. Chem. Soc.* **132,** 8787–8794 (2010).
63. Dall-Larsen, T., Kryvi, H. & Klungsoyr, L. Dinitrophenol, dicoumarol and pentachlorophenol as inhibitors and parasite substrates in the ATP phosphoribosyltransferase reaction. *Eur. J. Biochem.* **66,** 443–446 (1976).
64. Gohda, K. *et al.* QSAR Identification of Novel Potent Inhibitors for ATP-Phosphoribosyl Transferase Using Three-Dimensional Structural Database Search Technique. *Quant. Struct.-Act. Relat.* **20,** 143–147 (2001).
65. Cho, Y., Ioerger, T. R. & Sacchettini, J. C. Discovery of novel nitrobenzothiazole inhibitors for *Mycobacterium tuberculosis* ATP phosphoribosyl transferase (HisG) through virtual screening. *J. Med. Chem.* **51,** 5984–5992 (2008).
66. Mittlestead, G. Allosteric Regulation of the Adenosine Triphosphate Phosphoribosyltransferase from <i>Campylobacter jejuni</I>. (University of Canterbury, 31-

89, 2015).
67. Mittlestead, G. Allosteric Regulation of the Adenosine Triphosphate Phosphoribosyltransferase from *Campylobacter jejuni*. (University of Canterbury, 110-129, 2015).
68. Rao, S. T. & Rossmann, M. G. Comparison of super-secondary structures in proteins. *J. Mol. Biol.* **76,** 241–256 (1973).
69. Krissinel, E. & Henrick, K. Inference of Macromolecular Assemblies from Crystalline State. *J. Mol. Biol.* **372,** 774–797 (2007).
70. Pedreño, S., Pisco, J. P., Larrouy-Maumus, G., Kelly, G. & de Carvalho, L. P. S. Mechanism of feedback allosteric inhibition of ATP phosphoribosyltransferase. *Biochemistry* **51,** 8027–8038 (2012).
71. Lohkamp, B., Coggins, J. R. & Lapthorn, A. J. Purification , crystallization and preliminary X-ray crystallographic analysis of ATP-phosphoribosyl- transferase from *Escherichia coli* crystallization papers. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **D56,** 1488–1491 (2000).
72. Laidler, K. J. & King, M. C. The Development of Transition-State Theory. *J. Phys. Chem.* **87,** 2657–2664 (1983).
73. Eyring, H. The Activated Complex in Chemical Reactions. *J. Chem. Phys.* **3,** 107–115 (1935).
74. Truhlar, D. G., Garrett, B. C. & Klippenstein, S. J. Current Status of Transition-State Theory. *J. Phys. Chem.* **100,** 12771–12800 (1996).
75. Melander, L. & Saunders, W. H. in *Reaction rates of isotopic molecules* 1–3 (John Wiley & Sons, inc., 1980).
76. Kungwan, N. & Truong, T. N. Kinetics of the hydrogen abstraction $\cdot CH_3$ + alkane → CH_4 + alkyl reaction class: an application of the reaction class transition state theory. *J. Phys. Chem. A* **109,** 7742–7750 (2005).
77. Truhlar, D. G. & Garret, B. C. Variational transition state theory. *Ann. Rev. Phys. Chem* **35,** 159–189 (1984).
78. Pu, J. *et al.* Nonperfect synchronization of reaction center rehybridization in the transition state of the hydride transfer catalyzed by dihydrofolate reductase. *J. Am. Chem. Soc.* **127,** 14879–14886 (2005).
79. Singh, V. & Schramm, V. L. Transition-state structure of human 5'-methylthioadenosine phosphorylase. *J. Am. Chem. Soc.* **128,** 14691–14696 (2006).
80. Anslyn, E. & Dougherty, D. A. in *Modern physical organic chemistry* 355–482 (University Science Books, 2004). at <www.uscibooks.com>
81. International union of Pure and Applied Chemistry. International Union of Pure and Applied Chemistry Compendium of Chemical Terminology. *IUPAC* 1670 (2014).
82. Clayden, J., Greeves, N., Warren, S. & Wothers, P. in *Organic Chemistry* 105–110 (Oxford University Press, 2009).
83. Cleland, W. W. Use of Isotope Effects to Elucidate Enzyme Mechansims. *CRC Crit. rev. Biochem.* **13,** 385–428 (1982).
84. Schramm, V. L. Transition States, Analogues, and Drug Development. *ACS Chem. Biol.* **8,** 71–81 (2013).
85. Burgos, E. S., Vetticatt, M. J. & Schramm, V. L. Recycling nicotinamide. the transition-state structure of human nicotinamide phosphoribosyltransferase. *J. Am. Chem. Soc.* **135,** 3485–3493 (2013).
86. Scheuring, J. & Schramm, V. L. Kinetic isotope effect characterization of the transition state for oxidized nicotinamide adenine dinucleotide hydrolysis by pertussis toxin. *Biochemistry* **36,** 4526–4534 (1997).
87. Singh, V., Lee, J. E., Nunez, S., Howell, P. L. & Schramm, V. L. Transition State Structure of 5'-Methylthioadenosine/S-Adenosylhomocysteine Nucleosidase from *Escherichia coli* and Its Similarity to Transition State Analogues. *Biochem.* **44,** 11647–11659 (2005).
88. Cook, P. F. & Cleland, W. W. pH variation of isotope effects in enzyme-catalyzed reactions. 2.

Isotope-dependent step not pH dependent. Kinetic mechanism of alcohol dehydrogenase. *Biochemistry* **20,** 1805–1816 (1981).
89. Francis, K. & Kohen, A. Standards for the reporting of kinetic isotope effects in enzymology. *Perspect. Sci.* **1,** 110–120 (2014).
90. Roston, D. & Kohen, A. Elusive transition state of alcohol dehydrogenase unveiled. *Proc. Natl. Acad. Sci. U. S. A.* **107,** 9572–9577 (2010).
91. Cook, P. F. & Cleland, W. W. pH Variation of Isotope Effects in Enzyme-Catalyzed Reactions. 2. Isotope-Dependent Step Not pH Dependent. Kinetic Mechanism of Alcohol Dehydrogenase. *J. Am. Chem. Soc* **20,** 1805–1816 (1981).
92. Purich, D. L. in *Enzyme kinetics catalysis & control* 287–334 (Elsevier Academic, Elsevier Science, 2010).
93. Porter, D. J. & Spector, T. Alternative substrates for calf intestinal adenosine deaminase. A pre-steady-state kinetic analysis. *J. Biol. Chem.* **268,** 2480–2485 (1993).
94. Northrop, D. B. The expression of isotope effects on enzyme-catalyzed reactions. *Annu. Rev. Biochem.* **50,** 103–131 (1981).
95. Chuvikovsky, D. V *et al.* Ribokinase from E. coli: expression, purification, and substrate specificity. *Bioorg. Med. Chem.* **14,** 6327–32 (2006).
96. Parkin, D. W., Leung, H. B. & Schramm, V. L. Synthesis of Nucleotides with Specific Radiolabels in Ribose. *J. Biol. Chem.* **259,** 9411–9417 (1984).
97. Mo, H. & Raftery, D. Pre-SAT180, a Simple and Effective Method for Residual Water Suppression. *J. Magn. Reson.* **190,** 1–12 (2008).
98. Dimroth, O. Ueber intramolekulare Umlagerungen. Umlagerungen in der Reihe des 1, 2, 3-Triazols. *Justus Liebigs Ann. Chem.* **364,** 183–226 (1909).
99. Engel, J. d. Mechanism of the Dimroth Rearrangement in Adenosine. *Biochem. Biophys. Res. Commun.* **64,** 581–586 (1975).
100. Luo, M., Singh, V., Taylor, E. a & Schramm, V. L. Transition-state variation in human, bovine, and *Plasmodium falciparum* adenosine deaminases. *J. Am. Chem. Soc.* **129,** 8008–8017 (2007).
101. Apsel, B. *et al.* Targeted polypharmacology: discovery of dual inhibitors of tyrosine and phosphoinositide kinases. *Nat. Chem. Biol.* **4,** 691–699 (2008).
102. McCann, J. a B. & Berti, P. J. Transition state analysis of acid-catalyzed dAMP hydrolysis. *J. Am. Chem. Soc.* **129,** 7055–7064 (2007).
103. Milecki, J., Foldesi, A., Fischer, A., Adamiak, R. W. & Chattopadhyaya, J. Synthesis of multiply labelled ribonucleosides for sequence-specific labelling of oligo-RNA. *J. Label. Compd. Radiopharm.* **44,** 763–783 (2001).
104. Lewandowicz, A. & Schramm, V. L. Transition state analysis for human and Plasmodium falciparum purine nucleoside phosphorylases. *Biochemistry* **43,** 1458–1468 (2004).
105. Guthrie, R. D. & Jencks, W. P. IUPAC Recommendations for the Representation of Reaction Mechanisms. *Acc. Chem. Res.* **22,** 343–349 (1989).
106. Westheimer, F. H. The Magnitude of the Primary Kinetic Isotope Effect For Compounds of Hydrogen and Deuterium. *Chem. Rev.* **61,** 265–273 (1961).
107. Anisimov, V. & Paneth, P. ISOEFF98 . A program for studies of isotope effects using Hessian modifications. *J. Math. Chem.* **26,** 75–86 (1999).
108. Schramm, V. L. Enzymatic transition states, transition-state analogs, dynamics, thermodynamics, and lifetimes. *Ann. Rev. Biochem.* **80,** 703–732 (2011).
109. Alecu, I. M., Zheng, J., Zhao, Y. & Truhlar, D. G. Computational thermochemistry: Scale factor databases and scale factors for vibrational frequencies obtained from electronic model chemistries. *J. Chem. Theory Comput.* **6,** 2872–2887 (2010).
110. Scott, A. P. & Radom, L. Harmonic Vibrational Frequencies: An Evaluation of Hartree−Fock, Møller−Plesset, Quadratic Configuration Interaction, Density Functional Theory, and Semiempirical Scale Factors. *J. Phys. Chem.* **100,** 16502–16513 (1996).

111. Merrick, J. P., Moran, D. & Radom, L. An evaluation of harmonic vibrational frequency scale factors. *J. Phys. Chem. A* **111,** 11683–11700 (2007).
112. Schramm, V. L. & Shi, W. Atomic motion in enzymatic reaction coordinates. *Curr. Opin. Struct. Biol.* **11,** 657–65 (2001).
113. Frisch, M. J. *et al.* Gaussian 09, Revision D.01,. Gaussian, Inc., Wallingford CT, 2009.Gaussian 09 (2009).
114. Berti, P. J., Schramm, V. L., V, M. P. a, York, N. & April, R. V. Transition State Structure for the Hydrolysis of NAD+ Catalyzed by Diphtheria Toxin. *J. Am. Chem. Soc.* **119,** 12069–12078 (1997).
115. Rising, K. A. & Schramm, V. L. Transition State Analysis of NAD + Hydrolysis by the Cholera Toxin Catalytic Subunit. *J. Am. Chem. Soc.* **8,** 27–37 (1997).
116. Corfù, N. a & Sigel, H. Acid-base properties of nucleosides and nucleotides as a function of concentration. Comparison of the proton affinity of the nucleic base residues in the monomeric and self-associated, oligomeric 5'-triphosphates of inosine (ITP), guanosine (GTP), and ade. *Eur. J. Biochem.* **199,** 659–669 (1991).
117. Becke, A. D. Density - functional thermochemistry . III . The role of exact exchange. *J. Chem. Phys.* **98,** 5648–5652 (1993).
118. Zhao, Y. & Truhlar, D. G. The M06 suite of density functionals for main group thermochemistry, thermochemical kinetics, noncovalent interactions, excited states, and transition elements: Two new functionals and systematic testing of four M06-class functionals and 12 other function. *Theor. Chem. Acc.* **120,** 215–241 (2008).
119. Lee, C., Yang, W. & Parr, R. G. Into a Functional of the Electron Density F F. *Phys. Rev. B* **37,** 785–789 (1988).
120. Arunan, E. *et al.* Definition of the hydrogen bond (IUPAC Recommendations 2011). *Pure Appl. Chem.* **83,** 1637–1641 (2011).
121. Ehrlich, J. I. & Schramm, V. L. Electrostatic Potential Surface Analysis of the Transition State for AMP Nucleosidase and for Formycin 5'-Phosphate, a Transition-State Inhibitor. *Biochem.* **33,** 8890–8896 (1994).
122. Horenstein, B. & Schramm, V. Correlation of the molecular electrostatic potential surface of an enzymic transition state with novel transition-state inhibitors. *Biochemistry* **32,** 9917–9925 (1993).
123. Kline, P. C. & Schramm, V. L. Purine Nucleoside Phosphorylase. Catalytic Mechanism and Transition-State Analysis of the Arsenolysis Reaction. *biochem* **32,** 13212–13219 (1993).
124. Singh, V. & Schramm, V. L. Transition-State Analysis of *S. pneumoniae* 5'-Methylthioadenosine Nucleosidase. *J. Am. Chem. Soc.* **129,** 2783–2795 (2007).
125. Chen, X. Y., Berti, P. & Schramm, V. L. Transition-State Analysis for Depurination of DNA by Ricin A-Chain. *J Am Chem Soc* **122,** 6527–6534 (2000).
126. Singh, V., Luo, M., Brown, R. L., Norris, G. E. & Schramm, V. L. Transition-state structure of *Neisseria meningitides* 5 '-methylthioadenosine/S-adenosylhomocysteine nucleosidase. *J. Am. Chem. Soc.* **129,** 13831–13833 (2007).
127. Liao, R. Z. & Thiel, W. Determinants of regioselectivity and chemoselectivity in fosfomycin resistance protein FosA from QM/MM calculations. *J. Phys. Chem. B* **117,** 1326–1336 (2013).
128. Macrodomain, H., Hirsch, B. M., Burgos, E. S. & Schramm, V. L. Transition-State Analysis of 2-O-Acetyl-ADP-Ribose Hydrolysis by Human Macrodomain 1. *ACS Chem. Biol.* **9,** 2255–2262 (2014).
129. Ibarra, C., Nieslanik, B. S. & Atkins, W. M. Contribution of aromatic-aromatic interactions to the anomalous pK(a) of tyrosine-9 and the C-terminal dynamics of glutathione S-transferase A1-1. *Biochemistry* **40,** 10614–10624 (2001).
130. Versées, W., Loverix, S., Vandemeulebroucke, A., Geerlings, P. & Steyaert, J. Leaving group activation by aromatic stacking: an alternative to general acid catalysis. *J. Mol. Biol.* **338,** 1–6 (2004).

131. Garijo Añorbe, M. *et al.* Perturbation of the NH_2 pKa Value of Adenine in Platinum(II) Complexes: Distinct Stereochemical Internucleobase Effects. *Chem. - A Eur. J.* **10,** 1046–1057 (2004).
132. Olasz, A., Mignon, P., De Proft, F., Veszprémi, T. & Geerlings, P. Effect of the π-π stacking interaction on the acidity of phenol. *Chem. Phys. Lett.* **407,** 504–509 (2005).
133. Hammond, S. A Correlation of Reaction Rates. *J. Am. Chem. Soc.* **77,** 334–338 (1955).
134. Shi, W. *et al.* The 2.0 A structure of human hypoxanthine-guanine phosphoribosyltransferase in complex with a transition-state analog inhibitor. *Nat. Struct. Biol.* **6,** 588–593 (1999).
135. Tawada, Y., Tsuneda, T., Yanagisawa, S., Yanai, T. & Hirao, K. A long-range-corrected time-dependent density functional theory. *J. Chem. Phys.* **120,** 8425–8433 (2004).
136. Yanai, T., Tew, D. P. & Handy, N. C. A new hybrid exchange–correlation functional using the Coulomb-attenuating method (CAM-B3LYP). *Chem. Phys. Lett.* **393,** 51–57 (2004).
137. Heubel, P. H. C. & Popov, A. L. Acid Properties of Some Phosphonocarboxylic Acids. *J. Sol. Chem.* **8,** 615–625 (1979).
138. Mulliken, R. S. Electronic Population Analysis on LCAO-MO Molecular Wave Functions. IV. Bonding and Antibonding in LCAO and Valence-Bond Theories. *J. Chem. Phys.* **23,** 2343–2346 (1955).
139. Mulliken, R. S. Electronic Population Analysis on LCAO–MO Molecular Wave Functions. I. *J. Chem. Phys.* **23,** 1833–1840 (1955).
140. Lewis, B. E. & Schramm, V. L. Conformational Equilibrium Isotope Effects in Glucose by 13 C NMR Spectroscopy and Computational Studies †. *J. Am. Chem. Soc.* **123,** 1327–1336 (2001).
141. Taylor Ringia, E. a. *et al.* Transition state analogue discrimination by related purine nucleoside phosphorylases. *J. Am. Chem. Soc.* **128,** 7126–7127 (2006).
142. Luo, M., Li, L. & Schramm, V. L. Remote mutations alter transition-state structure of human purine nucleoside phosphorylase. *Biochemistry* **47,** 2565–2576 (2008).
143. Fedorov, a *et al.* Transition state structure of {PNP} and principles of atomic motion in enzymatic catalysis. *Biochemistry* **40,** 853–860 (2001).
144. Pauling, L. Molecular Architecture and Biological Reactions. *Chem. Eng. News* **24,** 1375–1377 (1946).
145. Wolfenden, R. Transition state analogues for enzyme catalysis. *Nature* **223,** 704–705 (1969).
146. Wolfenden, R. Analog approaches to the structure of the transition state in enzyme reactions. *Acc. Chem. Res.* **5,** 10–18 (1972).
147. Lienhard, G. E. Enzymatic catalysis and transition-state theory. *Science* **180,** 149–154 (1973).
148. Singh, V. *et al.* 5′-Methylthioadenosine/S-Adenosylhomocysteine Nucleosidase from *Escherichia coli* and Its Similarity to Transition State Analogues. *Biochemistry* **44,** 11647–11659 (2005).
149. Evans, G. B. *et al.* Second generation transition state analogue inhibitors of human 5′-methylthioadenosine phosphorylase. *J. Med. Chem.* **48,** 4679–4689 (2005).
150. Zhang, Y. *et al.* Transition state analogues of *plasmodium falciparum* and human orotate phosphoribosyltransferases. *J. Biol. Chem.* **288,** 34746–34754 (2013).
151. Mercer, T. B. *et al.* Looking glass inhibitors: both enantiomeric N-benzyl derivatives of 1,4-dideoxy-1,4-imino-d-lyxitol [a potent competitive inhibitor of α-d-galactosidase] and of 1,4-dideoxy-1,4-imino-l-lyxitol [a weak competitive inhibitor of α-d-galactosidase] inhibit na. *Tetrahedron: Asymmetry* **20,** 2368–2373 (2009).
152. Gregg, P. J., Kotuadsson, P., Oscarson, S. & Ruda, K. Synthesis of Part of a Proposed Insulin Second Messenger Glycosylinositol Phosphate and the Inner Core of Glycosylphosphatidylinositol Anchors. *Tetrahedron* **53,** 17727–17734 (1997).
153. Malladi, V. L. a, Sobczak, A. J., Meyer, T. M., Pei, D. & Wnuk, S. F. Inhibition of LuxS by S-ribosylhomocysteine analogues containing a [4-aza]ribose ring. *Bioorg. Med. Chem.* **19,** 5507–5519 (2011).
154. Waschke, D., Thimm, J. & Thiem, J. Highly Efficient Synthesis of Ketoheptoses. *Org. Lett.* **13,**

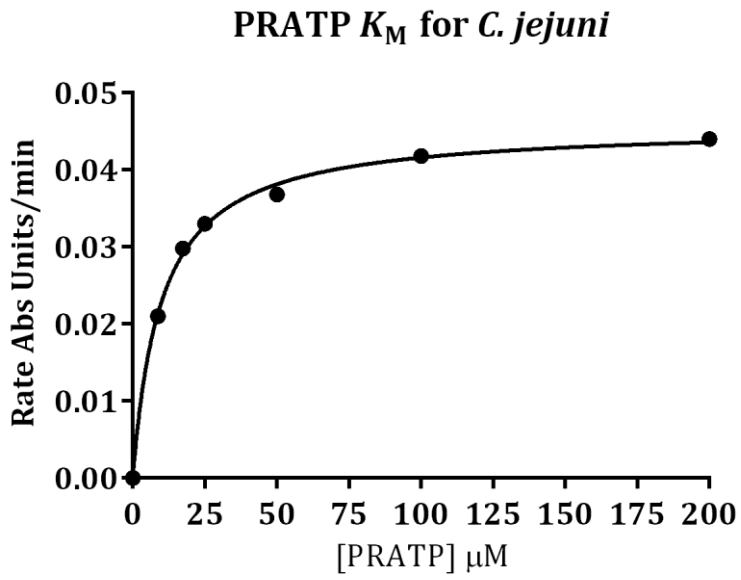
3628–3631 (2011).
155. Ohrui, H. & Emoto, S. Stereospecific synthesis of (+)-biotin. *Tetrahedron Lett.* **60,** 2765–2766 (1975).
156. Josan, J. S. & Eastwood, F. W. The introduction of unsaturated linkages into some carbohydrates by pyrolysis of cyclic, orthoformic esters. *Carb. Res.* **7,** 161–166 (1968).
157. Ovaa, H. *et al.* A flexible synthesis of cyclopentitol derivatives based on ring-closing metathesis of carbohydrate-derived 1,6-dienes. *J. Chem. Soc. Perkin Trans. 1* 2370–2377 (2002).
158. Schwab, P., Grubbs, R. H. & Ziller, J. W. Synthesis and Applications of $RuCl_2(=CHR')(PR_3)_2$: The Influence of the Alkylidene Moiety on Metathesis Activity. *J. Am. Chem. Soc.* **118,** 100–110 (1996).
159. Lohse-Fraefel, N. & Carreira, E. M. Polyketide building blocks via diastereoselective nitrile oxide cycloadditions with homoallylic alcohols and monoprotected homoallylic diols. *Chemistry* **15,** 12065–12081 (2009).
160. Baker, M. Fragment-based lead discovery grows up. *Nat. Rev. Drug. Discov.* **12,** 5–7 (2012).
161. Mashalidis, E. H., Śledź, P., Lang, S. & Abell, C. A three-stage biophysical screening cascade for fragment-based drug discovery. *Nat. Protoc.* **8,** 2309–2324 (2013).
162. May, P. C. *et al.* Robust central reduction of amyloid-β in humans with an orally available, non-peptidic β-secretase inhibitor. *J. Neurosci.* **31,** 16507–16516 (2011).
163. Jeppsson, F. *et al.* Discovery of AZD3839, a potent and selective BACE1 clinical candidate for the treatment of Alzheimers Disease. *J. Biol. Chem.* **287,** 41245–41257 (2012).
164. Kuntz, I. D., Chen, K., Sharp, K. a & Kollman, P. a. The maximal affinity of ligands. *Proc. Natl. Acad. Sci. U.S.A.* **96,** 9997–10002 (1999).
165. Hopkins, A. L., Groom, C. R. & Alex, A. Ligand efficiency: a useful metric for lead selection. *Drug Discov. Today* **9,** 430–431 (2004).
166. Rishton, G. M. Nonleadlikeness and leadlikeness in biochemical screening. *Drug Discov. Today* **8,** 86–96 (2003).
167. Lau, W. F. *et al.* Design of a multi-purpose fragment screening library using molecular complexity and orthogonal diversity metrics. *J. Comput. Aided Mol. Des.* **25,** 621–636 (2011).
168. Mpamhanga, C. P. *et al.* One scaffold, three binding modes: novel and selective Pteridine reductase 1 inhibitors derived from fragment hits discovered by virtual screening. *J. Med. Chem.* **52,** 4454–4465 (2009).
169. Congreve, M., Carr, R., Murray, C. & Jhoti, H. update discussion forum PET and knockout mice in A ' Rule of Three ' for fragment-based lead. *Drug Discov. Ther.* **8,** 876–877 (2003).
170. Meyer, B. & Peters, T. NMR Spectroscopy of Proteins NMR Spectroscopy Techniques for Screening and Identifying Ligand Binding to Protein Receptors Angewandte. *Angew. Chem. Int. Ed.* **42,** 864–890 (2003).
171. Chen, I.-J. & Hubbard, R. E. Lessons for fragment library design: analysis of output from multiple screening campaigns. *J. Comput. Aided Mol. Des.* **23,** 603–620 (2009).
172. Taylor, R. D., Jewsbury, P. J. & Essex, J. W. A review of protein-small molecule docking methods. *J. Comput. Aided. Mol. Des.* **16,** 151–166 (2002).
173. Jones, G., Willett, P., Glen, R. C., Leach, a R. & Taylor, R. Development and validation of a genetic algorithm for flexible docking. *J. Mol. Biol.* **267,** 727–748 (1997).
174. Chen, Y. & Shoichet, B. K. Molecular docking and ligand specificity in fragment-based inhibitor discovery. *Nat. Chem. Biol.* **5,** 358–364 (2009).
175. Kawatkar, S., Wang, H., Czerminski, R. & Joseph-McCarthy, D. Virtual fragment screening: an exploration of various docking and scoring protocols for fragments using Glide. *J. Comput. Aided. Mol. Des.* **23,** 527–539 (2009).
176. Zhao, H., Dong, J., Lafleur, K., Nevado, C. & Caflisch, A. Discovery of a Novel Chemotype of Tyrosine Kinase Inhibitors by Fragment-Based Docking and Molecular Dynamics. *ACS Med. Chem. Lett.* **3,** 834–838 (2012).

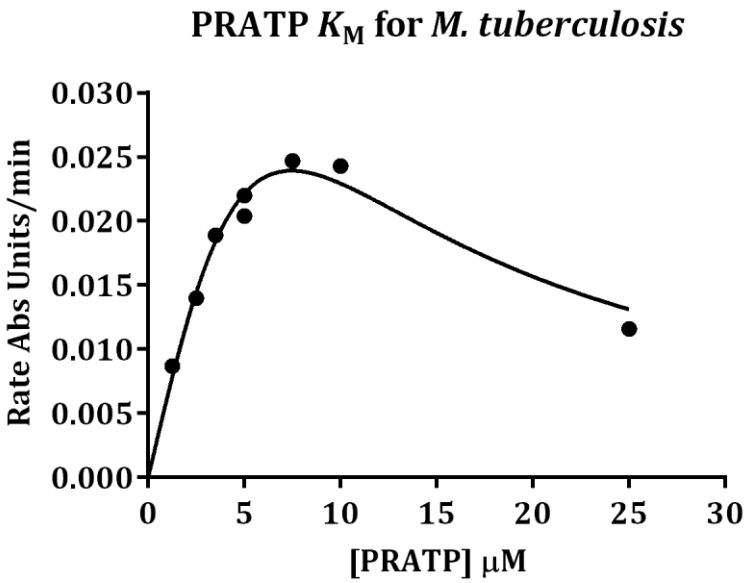
177. Good, A. C. *et al.* Implications of promiscuous Pim-1 kinase fragment inhibitor hydrophobic interactions for fragment-based drug design. *J. Med. Chem.* **55,** 2641–2648 (2012).
178. Wielens, J. *et al.* Parallel screening of low molecular weight fragment libraries: do differences in methodology affect hit identification? *J. Biomol. Screen.* **18,** 147–159 (2013).
179. Hubbard, R. E. Fragment approaches in structure-based drug discovery. *J. Synchrotron Radiat.* **15,** 227–30 (2008).
180. Friesner, R. a *et al.* Extra precision glide: docking and scoring incorporating a model of hydrophobic enclosure for protein-ligand complexes. *J. Med. Chem.* **49,** 6177–6196 (2006).
181. Friesner, R. a. *et al.* Glide: A New Approach for Rapid, Accurate Docking and Scoring. 1. Method and Assessment of Docking Accuracy. *J. Med. Chem.* **47,** 1739–1749 (2004).
182. Eldridge, M. D., Murray, C. W., Auton, T. R., Paolini, G. V & Mee, R. P. Empirical scoring functions: I. The development of a fast empirical scoring function to estimate the binding af nity of ligands in receptor complexes. *Proteus* **11,** 425–445 (1997).
183. Irwin, J. J., Sterling, T., Mysinger, M. M., Bolstad, E. S. & Coleman, R. G. ZINC: A Free Tool to Discover Chemistry for Biology. *J. Chem. Inf. Model.* **52,** 1757–1768 (2012).
184. Scott, D. E., Coyne, A. G., Hudson, S. a & Abell, C. Fragment-based approaches in drug discovery and chemical biology. *Biochem.* **51,** 4990–5003 (2012).
185. Mayer, M. & Meyer, B. Group epitope mapping by saturation transfer difference NMR to identify segments of a ligand in direct contact with a protein receptor. *J. Am. Chem. Soc.* **123,** 6108–17 (2001).
186. Celej, M. S., Montich, G. G. & Fidelio, G. D. Protein stability induced by ligand binding correlates with changes in protein flexibility. *Protein Sci.* **12,** 1496–1506 (2003).
187. Dai, R. *et al.* Fragment-Based Exploration of Binding Site Flexibility in *Mycobacterium tuberculosis* BioA. *J. Med. Chem.* **58,** 5208–5217 (2015).
188. Ferenczy, G. G. & Keseru, G. M. Thermodynamics of fragment binding. *J. Chem. Inf. Model.* **52,** 1039–1045 (2012).
189. Waldron, T. T. & Murphy, K. P. Stabilization of proteins by ligand binding: Application to drug screening and determination of unfolding energetics. *Biochemistry* **42,** 5058–5064 (2003).
190. Kawasaki, Y. & Freire, E. Finding a better path to drug selectivity. *Drug Discov. Today* **16,** 985–990 (2011).
191. Scott, A. D. *et al.* Thermodynamic optimisation in drug discovery: A case study using carbonic anhydrase inhibitors. *ChemMedChem* **4,** 1985–1989 (2009).
192. Larsson, E. A. *et al.* Fragment-based ligand design of novel potent inhibitors of tankyrases. *J. Med. Chem.* **56,** 4497–508 (2013).
193. Tébar, A. R. & Ballesteros, A. O. Kinetic Properties of ATP-Phosphoribosyltransferase of Escherichia coli. *Mol. Cell. Biochem.* **11,** 131–136 (1976).
194. Gee, C. T., Koleski, E. J. & Pomerantz, W. C. K. Fragment screening and druggability assessment for the CBP/p300 KIX domain through protein-observed 19F NMR spectroscopy. *Angew. Chemie - Int. Ed.* **54,** 3735–3739 (2015).
195. La, J. *et al.* Identification of mechanistically distinct inhibitors of HIV-1 reverse transcriptase through fragment screening. *Proc Natl Acad Sci U S A* **112,** 6979–6984 (2015).
196. Chimenti, M. S. *et al.* A Fragment-Based Ligand Screen Against Part of a Large Protein Machine: The ND1 Domains of the AAA+ ATPase p97/VCP. *J. Biomol. Screen.* **20,** 788–800 (2015).
197. Pangborn, A. B., Giardello, M. A., Grubbs, R. H., Rosen, R. K. & Timmers, F. J. Safe and Convenient Procedure for Solvent Purification. *Organometallics* **15,** 1518–1520 (1996).
198. Studier, F. W. Protein production by auto-induction in high density shaking cultures. *Protein Expr. Purif.* **41,** 207–234 (2005).
199. Kabsch, W. *XDS*. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **66,** 125–132 (2010).
200. Battye, T. G. G., Kontogiannis, L., Johnson, O., Powell, H. R. & Leslie, A. G. W. *iMOSFLM* : a new graphical interface for diffraction-image processing with *MOSFLM*. *Acta Crystallogr. Sect.*

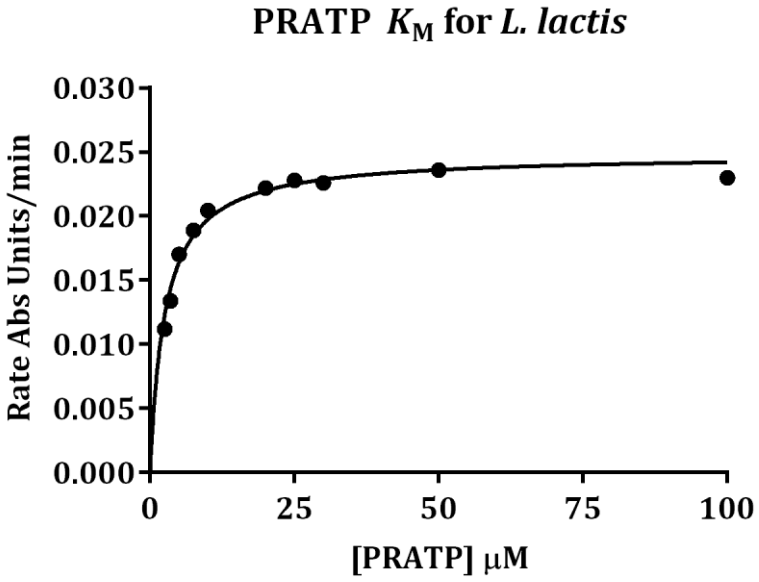
D Biol. Crystallogr. **67,** 271–281 (2011).
201. Winn, M. D. *et al.* Overview of the *CCP* 4 suite and current developments. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **67,** 235–242 (2011).
202. Evans, P. R. An introduction to data reduction: space-group determination, scaling and intensity statistics. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **67,** 282–292 (2011).
203. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr. D. Biol. Crystallogr.* **66,** 486–501 (2010).
204. Davis, I. W. *et al.* MolProbity: All-atom contacts and structure validation for proteins and nucleic acids. *Nucleic Acids Res.* **35,** 375–383 (2007).
205. Berman, H., Henrick, K. & Nakamura, H. Announcing the worldwide Protein Data Bank. *Nat. Struct. Biol.* **10,** 980 (2003).
206. Lanzetta, P. A., Alvarez, L. J., Reinach, P. S. & Candia, O. A. An Improved Assay for Nanomole Amounts of Inorganic Phosphate. *Anal. Biochem.* **100,** 95–97 (1979).
207. Dolomanov, O. V., Bourhis, L. J., Gildea, R. J., Howard, J. A. K. & Puschmann, H. OLEX2: A complete structure solution, refinement and analysis program. *J. Appl. Crystallogr.* **42,** 339–341 (2009).
208. Sheldrick, G. M. Crystal structure refinement with SHELXL. *Acta Crystallogr. Sect. C Struct. Chem.* **71,** 3–8 (2015).
209. Sheldrick, G. M. A short history of SHELX. *Acta Crystallogr. Sect. A Found. Crystallogr.* **64,** 112–122 (2007).
210. Ericsson, U. B., Hallberg, B. M., DeTitta, G. T., Dekker, N. & Nordlund, P. Thermofluor-based high-throughput stability optimization of proteins for structural studies. *Anal. Biochem.* **357,** 289–298 (2006).

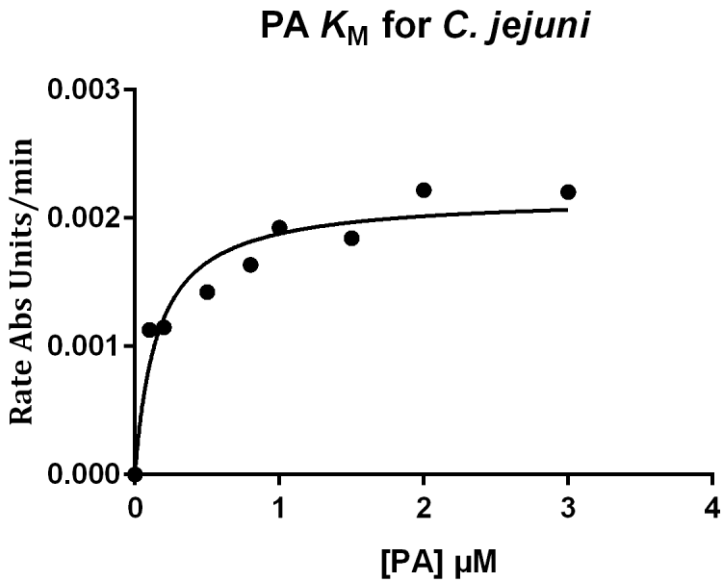
Appendix I

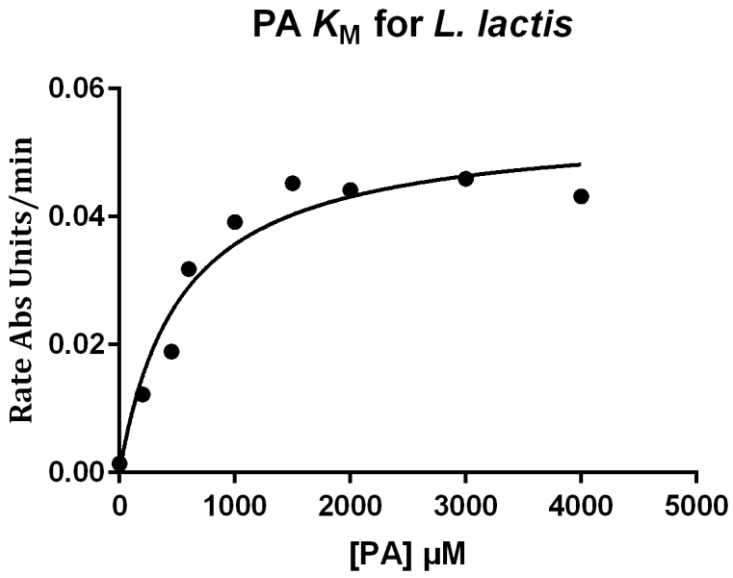
Michaelis-Menten plots for the determination of K_M values of PRATP, pyrophosphate and phosphonoacetic acid.

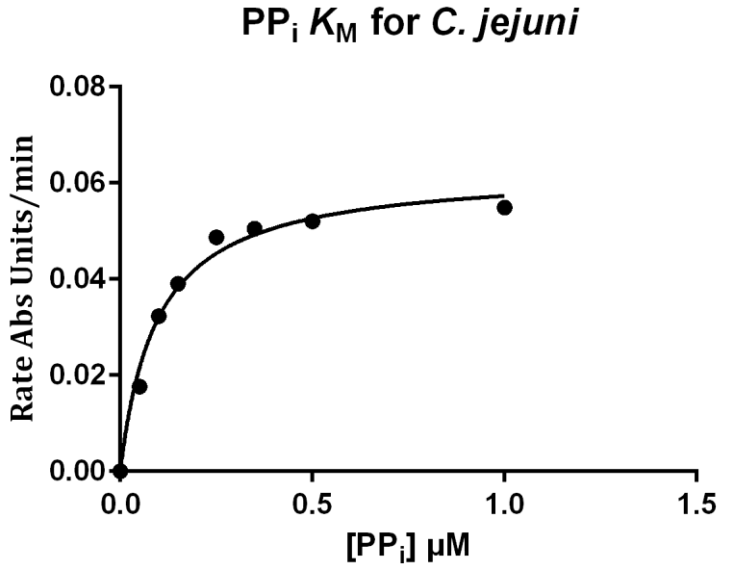


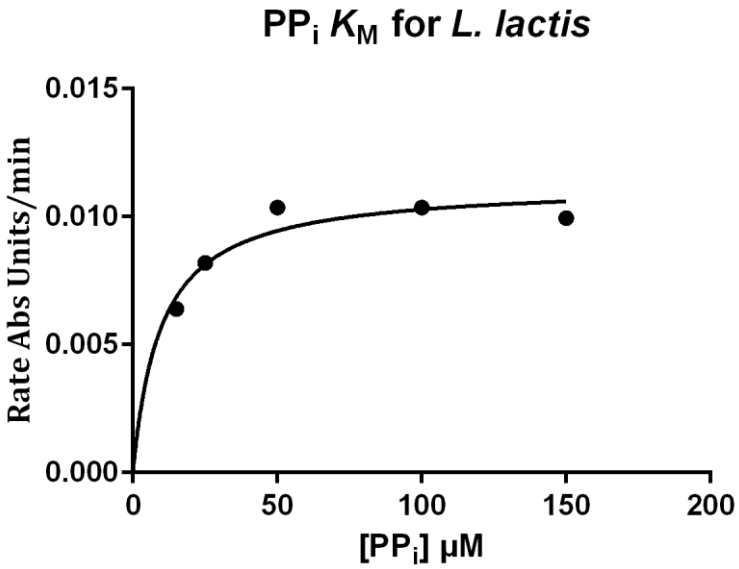



PRATP K_M for *L. lactis*
Rate Abs Units/min
0.030
0.025
0.020
0.015
0.010
0.005
0.000
0
25
50
75
100
[PRATP] μM


PA K_M for *C. jejuni*
Rate Abs Units/min
0.003
0.002
0.001
0.000
0
1
2
3
4
[PA] μM

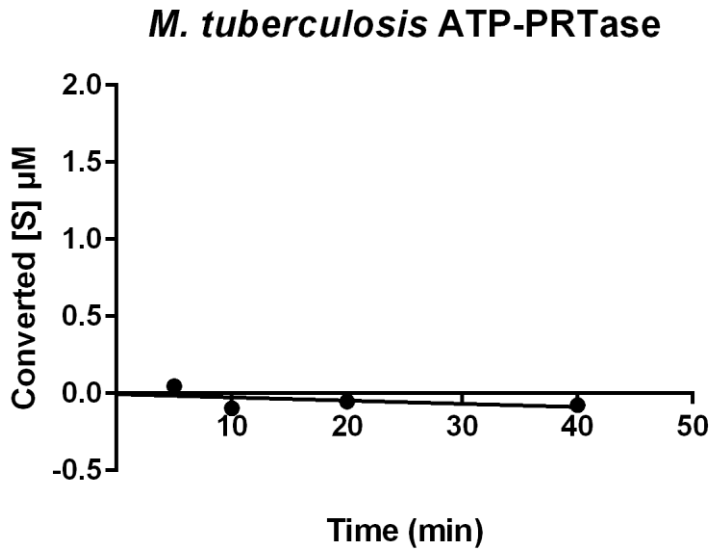

PA K_M for *L. lactis*
Rate Abs Units/min
0.06
0.04
0.02
0.00
0
1000
2000
3000
4000
5000
[PA] µM

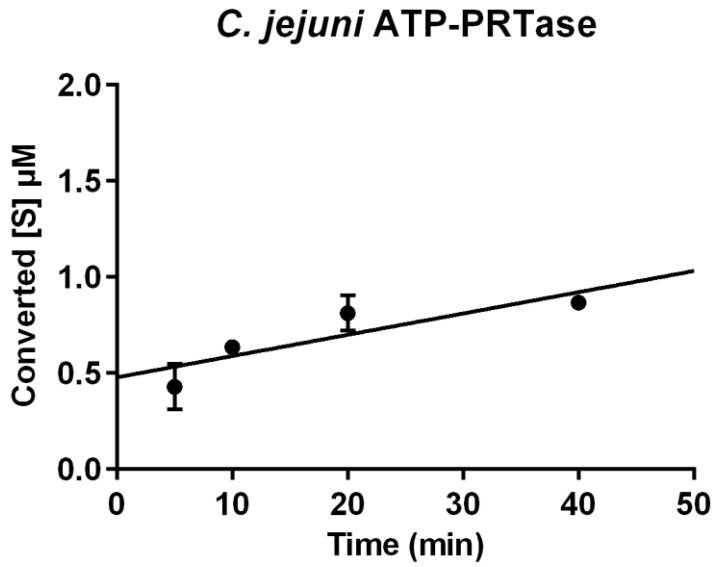

PP_i K_M for *C. jejuni*
Rate Abs Units/min
0.08
0.06
0.04
0.02
0.00
0.0
0.5
1.0
1.5
[PP_i] µM


PPi KM for L. lactis
Rate Abs Units/min
0.015
0.010
0.005
0.000
0
50
100
150
200
[PPi] µM

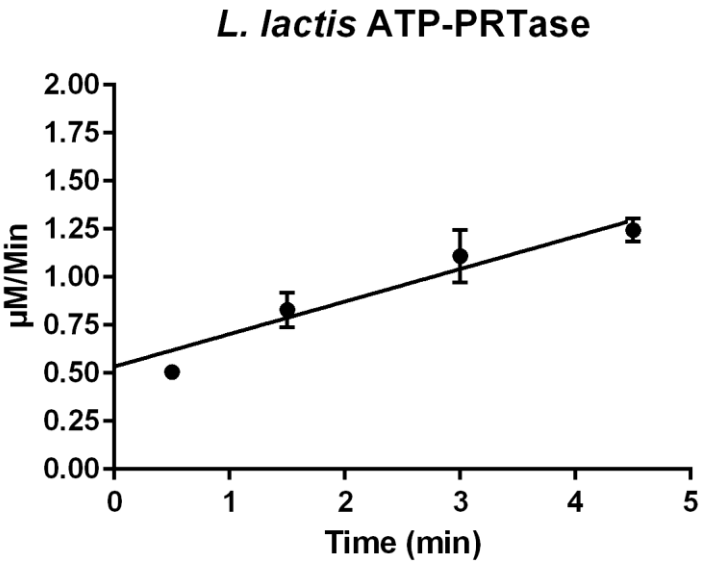
Appendix II

Plots for the determination of commitment factors





Y intercept = 0.480 µM



Y intercept = 0.530 µM

Kinetic isotope effect data for *M. tuberculosis* ATP-PRTase

[1′-^{3}H]	
KIE 1	1.225
KIE 2	1.245
KIE 3	1.217
KIE 4	1.283
KIE 5	1.249
KIE 6	1.259
KIE 7	1.265
KIE 8	1.259
Average	1.250
Stdev	0.021
SEM	0.008

[1′-^{14}C]	
KIE 1	1.027
KIE 2	1.033
KIE 3	1.024
KIE 4	1.026
KIE 5	1.027
KIE 6	1.028
KIE 7	1.027
KIE 8	1.030
Average	1.028
Stdev	0.003
SEM	0.001

[1-^{15}N]	
KIE 1	1.023
KIE 2	1.026
KIE 3	1.023
KIE 4	1.032
KIE 5	1.025
KIE 6	1.021
KIE 7	1.022
KIE 8	1.029
KIE 9	1.027
KIE 10	1.021
KIE 11	1.027
KIE 12	1.018
Average	1.024
Stdev	0.004
SEM	0.001

[6-^{15}N]	
KIE 1	1.011
KIE 2	1.016
KIE 3	1.014
KIE 4	1.020
KIE 5	1.010
KIE 6	1.011
KIE 7	1.009
KIE 8	1.010
Average	1.012
Stdev	0.004
SEM	0.001

[5′-^{3}H]	
KIE 1	1.014
KIE 2	1.022
KIE 3	1.020
KIE 4	1.021
KIE 5	1.014
KIE 6	1.017
KIE 7	1.011
KIE 8	1.018
Average	1.017
Stdev	0.004
SEM	0.001

Kinetic isotope effect data for *C. jejuni* ATP-PRTase

[1′-^{3}H]	
KIE 1	1.224
KIE 2	1.224
KIE 3	1.208
KIE 4	1.217
KIE 5	1.210
KIE 6	1.209
KIE 7	1.208
KIE 8	1.209
KIE 9	1.205
KIE 10	1.208
Average	1.212
Stdev	0.007
SEM	0.002

[1′-^{14}C]	
KIE 1	1.029
KIE 2	1.035
KIE 3	1.037
KIE 4	1.035
KIE 5	1.023
KIE 6	1.027
KIE 7	1.019
KIE 8	1.025
Average	1.029
Stdev	0.007
SEM	0.002

[1-^{15}N]	
KIE 1	1.027
KIE 2	1.028
KIE 3	1.030
KIE 4	1.031
KIE 5	1.029
KIE 6	1.028
KIE 7	1.024
KIE 8	1.025
Average	1.028
Stdev	0.002
SEM	0.001

[6-^{15}N]	
KIE 1	0.977
KIE 2	0.987
KIE 3	1.000
KIE 4	1.003
KIE 5	1.001
KIE 6	1.000
KIE 7	1.001
KIE 8	0.999
Average	0.996
Stdev	0.009
SEM	0.003

[5′-^{3}H]	
KIE 1	1.002
KIE 2	1.002
KIE 3	1.006
KIE 4	1.006
KIE 5	1.000
KIE 6	1.003
KIE 7	1.000
KIE 8	1.004
Average	1.002
Stdev	0.002
SEM	0.001

Kinetic isotope effect data for *L. lactis* ATP-PRTase

[1′-^{3}H]	
KIE 1	1.127
KIE 2	1.158
KIE 3	1.167
KIE 4	1.147
KIE 5	1.132
KIE 6	1.178
KIE 7	1.158
KIE 8	1.151
KIE 9	1.140
KIE 10	1.113
Average	1.147
Stdev	0.020
SEM	0.006

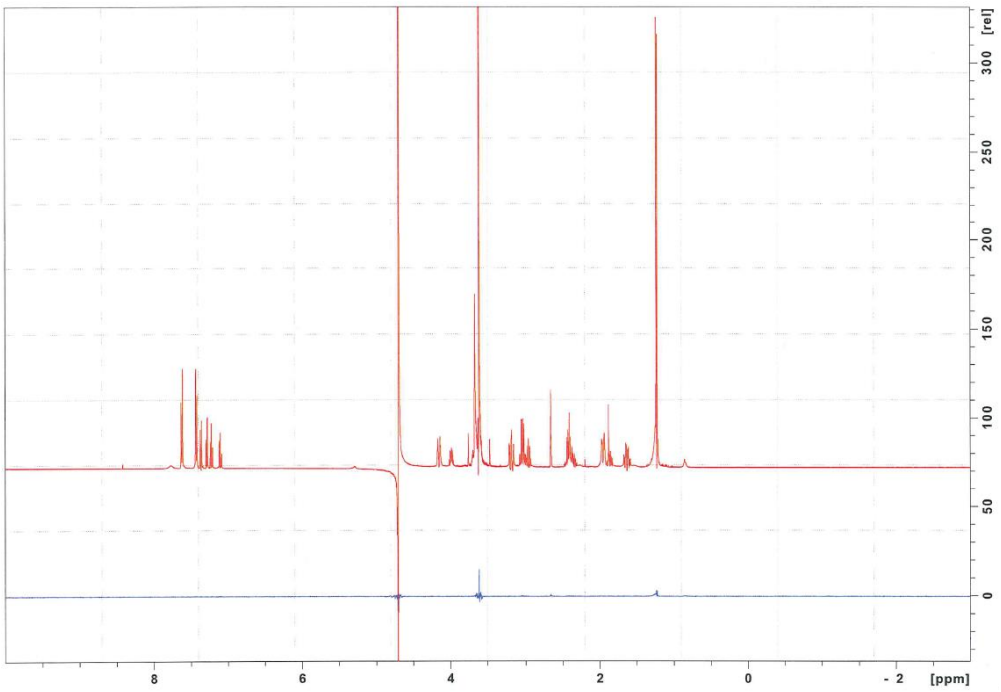
[1′-^{14}C]	
KIE 1	1.030
KIE 2	1.029
KIE 3	1.026
KIE 4	1.030
KIE 5	1.037
KIE 6	1.032
KIE 7	1.028
KIE 8	1.032
Average	1.031
Stdev	0.003
SEM	0.001

[1-^{15}N]	
KIE 1	0.981
KIE 2	0.985
KIE 3	0.994
KIE 4	0.998
KIE 5	0.999
KIE 6	1.003
KIE 7	0.999
KIE 8	1.003
Average	0.995
Stdev	0.008
SEM	0.003

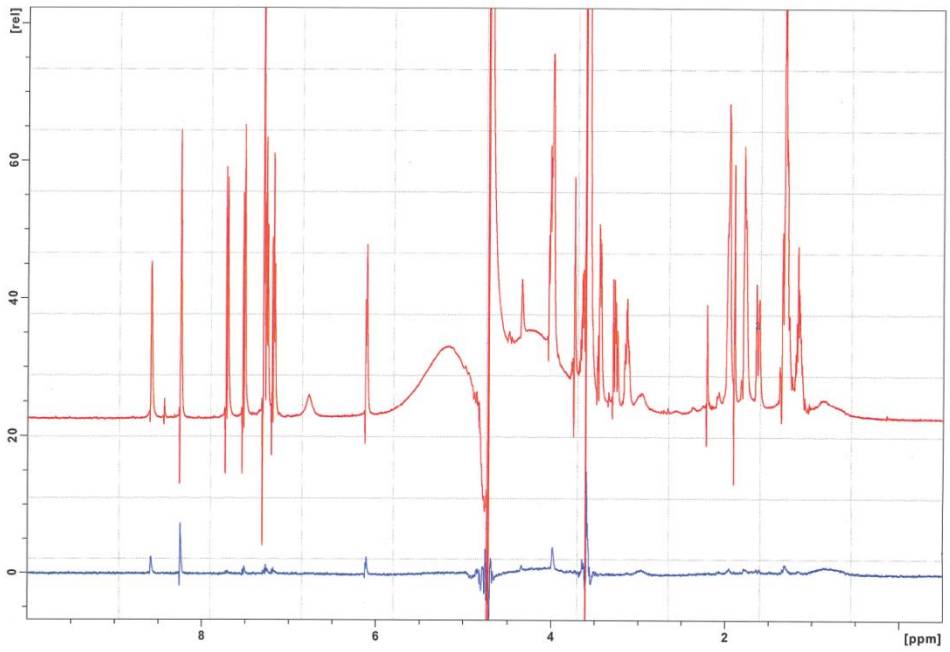
[6-^{15}N]	
KIE 1	1.006
KIE 2	1.007
KIE 3	1.006
KIE 4	1.004
KIE 5	1.002
KIE 6	1.003
KIE 7	1.002
KIE 8	1.008
Average	1.005
Stdev	0.002
SEM	0.001

[5′-^{3}H]	
KIE 1	1.012
KIE 2	1.008
KIE 3	1.013
KIE 4	1.010
KIE 5	1.003
KIE 6	1.008
KIE 7	1.009
KIE 8	1.006
Average	1.007
Stdev	0.003
SEM	0.001

Appendix III



In red is the reference spectrum. In blue the STD-NMR spectrum in absence of enzyme, no unspecific excitation is observed in the absence of the C. jejuni ATP-PRTase enzyme.



In red is the reference spectrum. In blue the STD-NMR spectrum in presence of the C. jejuni ATP-PRTase enzyme. The AMP ligand shows a stronger interaction then the other ligands.

Appendix IV

Crystal information for 5-*O*-*tert*-butyldimethylsilyl-2,3-*O*-isopropylidene-L-lyxono lactone 5.8

Crystal data and structure refinement for **5.8**.	
Identification code	GJM1a
Empirical formula	$C_{14}H_{26}O_5Si$
Formula weight	302.44
Temperature/K	120.01(10)
Crystal system	monoclinic
Space group	$P2_1$
a/Å	5.95513(12)
b/Å	10.14114(19)
c/Å	13.3855(3)
α/°	90
β/°	98.4789(19)
γ/°	90
Volume/Å^3	799.54(3)
Z	2
ρ_{calc}g/cm^3	1.256
μ/mm^{-1}	1.442
F(000)	328.0
Crystal size/mm^3	0.4068 × 0.3699 × 0.2228
Radiation	Cu Kα (λ = 1.5418)
2Θ range for data collection/°	13.376 to 147.52
Index ranges	$-7 \leq h \leq 7$, $-12 \leq k \leq 11$, $-16 \leq l \leq 12$
Reflections collected	7139
Independent reflections	3067 [R_{int} = 0.0193, R_{sigma} = 0.0195]
Data/restraints/parameters	3067/1/188
Goodness-of-fit on F^2	1.034
Final R indexes [I>=2σ (I)]	R_1 = 0.0253, wR_2 = 0.0689
Final R indexes [all data]	R_1 = 0.0255, wR_2 = 0.0690
Largest diff. peak/hole / e Å^{-3}	0.25/-0.17
Flack parameter	0.01(3)